Metodología de análisis de contenido

Teoría y práctica

Klaus Krippendorff

Paidós Comunicación



Klaus Krippendorff

Metodología de análisis de contenido

Teoría y práctica



Título original: *Content Analysis. An Introduction to its Methodology*
Publicado en inglés por Sage Publications, Inc., Newbury Park

Traducción de Leandro Wolfson
Supervisión de José Manuel Pérez Tornero

Cubierta de Mario Eskenazi

1.ª edición, 1990
1.ª reimpresión, 1997

ISBN: 84-7509-627-1
Depósito legal: B-28.603/1997

Impreso en Hurope, S. L.,
Recaredo, 2 - 08005 Barcelona

Impreso en España - Printed in Spain

Sumario

Prefacio

El análisis de contenido puede llegar a convertirse en una de las más importantes técnicas de investigación de las ciencias sociales. Procura comprender los datos, no como un conjunto de acontecimientos físicos, sino como fenómenos simbólicos, y abordar su análisis directo. Los métodos de las ciencias naturales no se ocupan necesariamente de significados, referencias, valoraciones e intenciones. Los métodos de la investigación social derivados de esas "rigurosas" disciplinas se las ingenian para dejar de lado, en aras de la conveniencia, esos fenómenos. Sin embargo, nadie puede poner en duda la importancia que tienen los símbolos en la sociedad.

En la actualidad, el análisis de contenido se encuentra, en cierto sentido, en una encrucijada. Sus raíces están en la fascinación periodística por los números, que supuestamente consigue una enunciación cuantitativa más convincente que una cualitativa. El análisis de contenido podría continuar con este juego de cómpu-

tos que quizá sea excitante, pero que no conduce a la comprensión. También podría emprender, con más seriedad que en el pasado, lo que corresponde a la tarea de analizar algo tan obviamente simbólico como un fenómeno simbólico, reconociendo su papel social, sus efectos y su significado. El empleo de los ordenadores en el procesamiento del lenguaje natural nos ha enseñado con meridiana claridad que esto requiere mucha más imaginación y trabajo teórico de lo que se suponía. También requiere una lógica apropiada y una metodología que se adecue a esta finalidad.

El presente libro está dirigido a un público muy amplio de profesionales de las ciencias sociales, investigadores de las comunicaciones y estudiantes. En él se desarrolla una nueva perspectiva. Puede servir como texto en diversos planes de estudio de ciencias sociales. Puede actuar como guía práctica en campos de investigación que incluyan fenómenos simbólicos. Y también proporciona material que ha de ser útil para el usuario que quiera evaluar críticamente los hallazgos del análisis del contenido.

Debo dar las gracias a numerosos estudiantes de la Annenberg School of Communication, de Filadelfia, ya que de nuestra mutua interacción surgió la organización conceptual de este libro; también quiero expresar mi gratitud a los diversos proyectos de investigación de análisis del contenido cuyos problemas conceptuales y metodológicos, por las repeticiones y pautas que evidenciaban, merecían soluciones más generales; por último, deseo expresar mi agradecimiento al Instituto de Estudios Avanzados de los Países Bajos, por haberme proporcionado el lugar en el que fue posible completar este trabajo.

Klaus Krippendorff
Universidad de Pennsylvania

Introducción

La expresión "análisis del contenido" tiene unos cincuenta años de antigüedad. Si bien el *Webster's Dictionary of the English Language* sólo la incluye desde el año 1961, sus orígenes intelectuales se remontan muy lejos en la historia, hasta el inicio del uso consciente de los símbolos y del lenguaje por parte del ser humano. Esa conciencia, que reemplazó a la magia, no sólo se expresó a través de diversas disciplinas antiguas (la filosofía, la retórica, el arte, la criptografía), sino que provocó además inquisiciones religiosas y censuras políticas por parte de monarcas y gobernantes. Hoy, el interés por los fenómenos simbólicos ya está institucionalizado en la literatura, la crítica de los medios de comunicación de masas y la educación, así como en disciplinas académicas como la antropología, la lingüística, la psicología social, la sociología del conocimiento, y en numerosas actividades profesionales o prácticas, como la psicoterapia, la publicidad, la política, etc. Prácticamente toda la gama de las humanidades y

de las ciencias sociales, incluidos los intentos de mejorar las condiciones políticas y sociales de vida, se ocupan de símbolos, significados y mensajes, de sus funciones y de sus efectos.

Pero, por antiguas que sean estas raíces, su existencia, ya que no su persistencia, no debe hacernos olvidar que los esfuerzos modernos realizados en este sentido son significativamente distintos, tanto en su método como en sus finalidades. A nuestro juicio, poseen tres características diferenciales.

En primer lugar, *el análisis de contenido tiene una orientación fundamentalmente empírica*, exploratoria, vinculada a fenómenos reales y de finalidad predictiva. Si bien muchos de los términos actuales relacionados con el lenguaje tienen origen griego (por ejemplo "signos", "significado", "símbolos" y "lógica"), el interés de los griegos por el lenguaje era en gran medida prescriptivo y clasificatorio. La lógica aristotélica tenía como propósito dar origen a expresiones claras, y gran parte de la retórica estaba dedicada a una tipología de las argumentaciones apropiadas. Sólo en los últimos tiempos se han traspasado las barreras del pensamiento normativo. Por ejemplo, los estudios sobre el comportamiento inteligente nos han indicado que el cerebro humano no se amolda en líneas generales a las categorías aristotélicas. Poco a poco parece haber ocupado su sitio una "psicología" no aristotélica.

Otros cambios semejantes hacia una orientación abiertamente empírica distinguen el análisis de contenido de las epistemologías clásicas. Si bien son indispensables las especulaciones, generalizaciones y construcciones teóricas, en última instancia éstas deben demostrar sus méritos cuando se aplican a los datos reales. Con esta orientación empírica, el análisis de contenido se ha sumado a otros métodos de investigación que contribuyen al conocimiento, especializándose sin embargo en hechos simbólicos frente a los cuales los restantes métodos son, por lo general, insensibles.

En segundo lugar, *el análisis de contenido trasciende las nociones convencionales del contenido* como objeto de estudio, y está estrechamente ligado a concepciones más recientes sobre los fenómenos simbólicos. Esto puede situarse dentro del contexto de una conciencia distinta acerca de la comunicación humana, de los nuevos medios de comunicación y del papel que éstos desempeñan en la transmisión de información dentro de la sociedad. Es probable que hoy nos encontremos inmersos en la última

de cuatro revoluciones concernientes a los conceptos de la comunicación, que han estado relacionadas con las siguientes ideas:

- La idea de *mensajes* (la conciencia de la naturaleza simbólico-representativa de los intercambios humanos) tiene sus orígenes en una fecunda combinación del comercio intercultural y de la ciencia, que hizo su aparición en la antigua Grecia.
- La idea de *canales* (la conciencia de las limitaciones que impone a la expresión humana la elección de un determinado medio) tiene sus orígenes en el empleo cada vez mayor de la tecnología de las comunicaciones a partir de la imprenta y, más tarde, del uso de medios electrónicos.
- La idea de *comunicación* (la conciencia de las dependencias interpersonales, las relaciones sociales, la estructura y la estratificación social que crea, de modo subrepticio, el intercambio de información) tiene sus orígenes en los veloces cambios sociales, incluida la decadencia de las instituciones sociales y de las relaciones humanas tradicionales, desde comienzos del presente siglo.
- La idea de *sistema* (la conciencia de las interdependencias globales y dinámicas) tiene sus orígenes en el difundido uso de la compleja tecnología de las comunicaciones, de los medios de comunicación de masas, de las redes de transmisión de canales múltiples y los ordenadores, con la consecuente dispersión de las formas organizativas y el entrecruzamiento de las empresas industriales privadas, los organismos gubernamentales, los medios de comunicación de masas y otras instituciones.

De todo ello se desprende que el interés empírico por los hechos simbólicos ya no puede aplicarse al estudio de los mensajes de forma aislada, ni reducir la comunicación a un proceso psicológico o considerar las interpretaciones lingüísticas de un mensaje como la base de la explicación. Los cambios producidos en la trama social exigen una definición *estructural* del contenido, que tenga en cuenta los canales y las limitaciones de los flujos de información, los procesos de comunicación, y sus funciones y efectos en la sociedad, los sistemas que incluyen tecnología avanzada y las modernas instituciones sociales. Debemos admitir esto, en una posible definición del análisis de contenido, no ya por su ámbito tradicional de aplicación (el significado de los mensajes), sino por el proceso implicado en el análisis de los datos como entidades simbólicas (y no en el análisis de estas entidades por sí mismas). Quizá la expresión "análisis de conte-

nido" ya no resulte apropiada para este contexto más amplio, en cuyo interior se entienden (y deben entenderse) los mensajes y los datos simbólicos.

En tercer lugar, *el análisis de contenido está desarrollando una metodología propia* que permite al investigador programar, comunicar y evaluar críticamente un plan de investigación con independencia de sus resultados. La necesidad de una metodología de esta índole se pone claramente de manifiesto a causa de la amplia perspectiva que ha adoptado el análisis de contenido, la incapacidad cada vez mayor de un individuo aislado para tomar conciencia de lo que pueden representar o producir los sucesos simbólicos complejos, y el hecho de apoyarse cada vez más en un tipo de investigación organizada que exige coordinar los controles de calidad y utilizar técnicas fiables.

Dado que precisamente de esto trata este libro, la metodología no debe confundirse con la materia de la que se ocupa. El propósito de la metodología es describir y examinar la lógica de la creación de métodos y técnicas de investigación, poner de relieve su eficacia y sus limitaciones, generalizar sus éxitos y fracasos, descubrir los ámbitos apropiados de aplicación y predecir sus posibles contribuciones al conocimiento. El desarrollo de una metodología constituye un paso fundamental en la historia natural de cualquier empresa científica. Por ejemplo, durante milenios se narraron relatos o historias antes de que von Ranke, hace ya alrededor de un siglo, diera al "documento" el estatuto metodológico que hoy posee en el estudio de la historia. También el análisis de contenido fue práctica antes de que se codificara. Aunque puede aducirse que todos los análisis de contenido son diferentes entre sí, y que cada disciplina que emplea esta técnica aborda problemas algo distintos, debemos decir que todos los análisis de contenido comparten una lógica de composición, una forma de razonamiento y ciertos criterios de validez. En estos procedimientos comunes radica, a nuestro juicio, la materia de la que se ocupa la metodología del análisis del contenido.

Discrepamos con la afirmación de que "el análisis de contenido no es otra cosa que lo que todo el mundo hace al leer un periódico, pero a mucha mayor escala". Quizá el análisis de contenido se inició de esta manera, y su metodología no impida que se emplee con esa orientación; pero la caracterización anterior remite a una etapa periodística, ya superada, del análisis de contenido. En nuestra condición de lectores de periódicos, tenemos todo el

derecho a utilizar nuestra cosmovisión peculiar y de recurrir a nuestras preferencias; pero como investigadores lo mejor es que evitemos toda tendenciosidad, desconfiemos de la interpretación de un único individuo, expliquemos lo que estamos haciendo, comuniquemos nuestros hallazgos para que otros puedan examinarlos o reproducirlos y, por encima de todo, tomemos conciencia de la diferencia cualitativa existente entre una metodología que nos proporciona una plataforma a partir de la cual podemos referirnos a los datos y los procedimientos científicos y, por otro lado, lo que esos fenómenos significan individualmente para cada uno de nosotros.

Este libro constituye una introducción a la manera de analizar los datos como comunicaciones simbólicas. Sus capítulos pueden agruparse en tres grandes categorías. La parte teórica (capítulos 1, 2 y 3) se inicia con una breve historia del análisis de contenido, establece una definición que lo distingue de otros métodos, y ejemplifica su ámbito de aplicación práctica. La parte metodológica-práctica (capítulos 4 al 11) comienza con la lógica de los proyectos de análisis de contenido, para luego ocuparse de los procedimientos utilizados: determinación de unidades, muestreo, registro, construcción de lenguajes de datos, construcciones analíticas, técnicas de cálculo y uso del ordenador. La parte crítica (capítulos 12 y 13) se interesa por los dos principales criterios de calidad de los análisis de contenido: la fiabilidad y la validez. El libro concluye con una guía práctica en la que se sintetiza lo anterior desde esta particular perspectiva.

Quien jamás haya practicado un análisis de contenido puede que desee familiarizarse primero con los fundamentos conceptuales, para luego iniciar un pequeño proyecto propio, tal vez siguiendo los pasos que se indican en el capítulo 14. Al tratar de describir y justificar lo que hace cuando interpreta sus datos simbólicamente, pronto advertirá todo lo que ello implica desde el punto de vista del procedimiento, así como los conocimientos sustantivos acerca de las condiciones contextuales de sus datos que necesita para lanzarse a diseñar un proyecto más amplio. Estos conocimientos sustantivos son propios de cada disciplina o se refieren a un objeto de análisis concreto, y deberán adquirirse en otra parte. El investigador que ya ha realizado alguna vez un análisis de contenido podrá obtener en este libro una perspectiva más amplia y nuevas herramientas para el análisis, así como nuevas ideas acerca de su metodología, y quizás asegurarse de que

sus hallazgos puedan defenderse a la luz del marco de referencia conceptual que aquí se desarrolla. El usuario crítico de los resultados de análisis de contenido efectuados en otros lugares podrá aprovechar al máximo las técnicas y criterios de evaluación, ya que éstos le permitirán valorar hasta qué punto puede fiarse de esos resultados, y le harán adoptar una actitud crítica.

Este libro habrá alcanzado sus propósitos si puede contribuir a mejorar la consideración social de alguno de los análisis de contenido que, en número cada vez mayor, se están llevando a cabo en el campo de las humanidades y de las ciencias sociales, y si estimula el desarrollo de posteriores métodos de indagación de la realidad simbólica.

1. Historia

Las indagaciones empíricas sobre el contenido de las comunicaciones se remontan a los estudios teológicos de fines del siglo XVII, cuando la Iglesia estaba inquieta por la difusión de los temas de índole no religiosa a través de los periódicos. A partir de entonces fueron ganando terreno en numerosas esferas. En este capítulo se pasa revista al desarrollo histórico de la metodología del análisis de contenido, haciendo referencia a los estudios a gran escala de la prensa, a las investigaciones sociológicas y lingüísticas, y en especial a los estudios de los nuevos medios de comunicación, efectuados con diferentes propósitos, desde los que se ocupan de los símbolos y la propaganda política, hasta los que abordan los mitos, las leyendas populares y los acertijos.

El primer caso bien documentado de análisis cuantitativo de material impreso tuvo lugar en Suecia en el siglo XVIII. DOVRING (1954-1955) describió este episodio, incluido en una colección de noventa himnos de autor desconocido, titulada *Los cantos de Sion.* Esta colección de himnos logró pasar la censura oficial, pero pronto se la acusó de socavar la moral del clero ortodoxo de la Iglesia oficial sueca, alegando que sus cánticos eran "populares y contagiosos" y contribuían a la exaltación de un grupo disidente. Un hecho destacado fue la participación de grandes eruditos en esta controversia, que cristalizó en torno a la cuestión de si los cánticos eran realmente portadores de ideas perniciosas. Uno de los bandos comenzó a realizar el cómputo de los símbolos religiosos que aparecían en ellos, mientras otro contaba esos mismos símbolos en el cantoral canónico, y no pudieron encontrar ninguna diferencia. Luego se computaron los símbolos en los contextos en que aparecían,

comparándolos con el resultado de un estudio alemán acerca de los hermanos moravos, una secta que había sido declarada ilegal, hasta que se puso de manifiesto en qué radicaba la diferencia y cómo se podía explicar. Para estos estudiosos, la polémica estimuló un debate metodológico, y la cuestión se zanjó finalmente en este ámbito.

LOEBL (1903) publicó en alemán un elaborado esquema clasificatorio para el análisis de la "estructura interna del contenido" de acuerdo con las funciones sociales que desempeñan los periódicos. Su libro adquirió celebridad en los círculos periodísticos, pero no estimuló investigaciones empíricas posteriores.

En la primera reunión de la Sociedad Sociológica Alemana, en 1910, Max WEBER (1911) propuso llevar a cabo un amplio análisis de contenido de los medios de prensa, pero, por una variedad de razones, el estudio no pudo llevarse a la práctica. Por esa época, MARKOV (1913) trabajaba en una teoría sobre las cadenas de símbolos, y publicó un análisis estadístico de una muestra extraída de la novela en verso de Pushkin, *Eugenio Onegin*. En su mayoría, estas investigaciones fueron descubiertas sólo muy recientemente, o bien influyeron de manera indirecta en la bibliografía sobre el análisis de contenido. Por ejemplo, el trabajo de Markov ingresó en esta bibliografía por mediación de Shannon (SHANNON y WEAVER, 1949) con su teoría de la información, así como a través del subsiguiente análisis de la contingencia de OSGOOD (1959).

ANALISIS CUANTITATIVO DE PERIODICOS

Hacia fines del siglo pasado se produjo un visible aumento de la producción masiva de material impreso en los Estados Unidos, así como de la inquietud por evaluar los mercados de masas y conocer la opinión pública. Surgieron escuelas de periodismo que plantearon la exigencia de que se cumplieran ciertas normas éticas y de que se efectuaran investigaciones empíricas acerca del fenómeno del periódico moderno. Estas demandas, sumadas a una noción algo simplista de la objetividad científica, quedaron satisfechas por lo que entonces se denominó *análisis cuantitativo de periódicos*.

El que fue probablemente uno de los primeros análisis de esta índole, publicado en 1893, formulaba esta intencionada pregunta:

"Do newspapers now give the news?"* (SPEED, 1893). El autor mostraba en este artículo el modo en que las cuestiones religiosas, científicas y literarias habían desaparecido de los principales periódicos neoyorkinos entre 1881 y 1893, para dejar lugar a la chismografía, los escándalos y los deportes. Un estudio similar trató de poner de manifiesto el abrumador espacio destinado a asuntos "desmoralizadores", "malsanos" y "triviales", en oposición a los temas "que valían la pena" (MATHEWS, 1910). Midiendo simplemente los centímetros de columna que un periódico destinaba a determinados temas, los periodistas trataron de revelar "la verdad acerca de los periódicos" (STREET, 1909). Creían que habían encontrado la forma de poner de relieve que la búsqueda del lucro era la causa del "chabacano periodismo amarillo" (WILCOX, 1900), estaban convencidos de que habían logrado corroborar "la influencia de ciertos tratamientos periodísticos sobre el aumento del delito y otras actividades antisociales" (FENTON, 1910), e incluso llegaron a la conclusión de que "el estudio del contenido de la prensa a lo largo de un cuarto de siglo demostraba la existencia de una creciente demanda de hechos" (WHITE, 1924).

El problema metodológico dominante parecía ser el apoyo que los datos científicos ofrecían a las argumentaciones periodísticas. Para ser irrefutables, dichos datos tenían que ser "cuantitativos". La veneración por los números es, sin duda, antigua. En una nota a pie de página, BERELSON y LAZARSFELD (1948, pág. 9) citan una fuente de casi doscientos años atrás:

> Quizá donde mejor se refleje el espíritu de esta batalla por la ratificación sea en el credo que, irónicamente, cada uno de los bandos contendientes atribuía a su oponente. La receta de un ensayo "antifederalista", que indica de manera muy sintética el tipo de orientación que movía a quienes se oponían a la Constitución norteamericana, rezaba así: "Bien nacido, nueve veces; aristocracia, dieciocho veces; libertad de prensa, repetida trece veces; libertad de conciencia, una vez; esclavitud de los negros, una vez; juicio con jurado, siete veces; grandes hombres, repetido seis veces; mister Wilson, cuarenta veces... mezclen todos estos ingredientes y sírvanlos a su gusto" (extraído de *New Hampshire Spy*, 30 de noviembre de 1787).

* ¿Se ocupan los periódicos de ofrecer noticias? [E.]

No obstante, el análisis cuantitativo de los periódicos aportó muchas ideas valiosas. En 1912, Tenney planteó la necesidad de una encuesta permanente y a gran escala acerca del contenido de los medios de prensa, a fin de establecer un sistema de control del "clima social" que pudiera compararse en precisión con las estadísticas del Servicio Meteorológico de los Estados Unidos. Lo demostró con algunos periódicos neoyorkinos destinados a los distintos grupos étnicos, pero entonces no era factible llevar a cabo aplicaciones prácticas. Este enfoque del análisis del contenido llegó a su culminación con el estudio del sociólogo WILLEY titulado *The Country Newspaper* [El periódico rural] (1926). En este paradigmático estudio, Willey trazó los orígenes de los semanarios rurales del estado de Connecticut, dando cuenta de su tirada, los cambios acaecidos en su temática y el papel social que fueron adquiriendo en su competencia con los periódicos de las grandes ciudades.

Al cobrar importancia otros medios de comunicación de masas, este método consistente en medir la cantidad de material impreso clasificándolo por categorías temáticas se extendió inicialmente a la radio (ALBIG, 1938) y más tarde al cine y la televisión. Este tipo de análisis continúa aplicándose aún hoy a una amplia variedad de contenidos, como los de los libros de texto, las historietas, los discursos pronunciados en público y la publicidad.

LOS PRIMEROS ANALISIS DE CONTENIDO

La segunda fase del desarrollo intelectural del análisis de contenido fue la consecuencia, como mínimo, de tres factores. En primer término, los nuevos y poderosos medios electrónicos de comunicación ya no podían considerarse una mera extensión del periódico. En segundo lugar, en el período posterior a la crisis económica surgieron numerosos problemas sociales y políticos respecto de los cuales se suponía que los nuevos medios de comunicación de masas habían desempeñado un papel causal. En tercer lugar, debe mencionarse la aparición de los métodos empíricos de investigación en las ciencias sociales.

Por ejemplo, la sociología comenzó a hacer amplio uso de las investigaciones mediante encuestas y de los sondeos de opinión. Las experiencias recogidas en el análisis de la opinión pública dieron origen al primer examen serio de los problemas metodológicos del análisis de contenido, llevado a cabo por WOODWARD y

titulado "Quantitative Newspaper Analysis as a Technique of Opinion Research" (1934). A partir de los escritos sobre la opinión pública, se incorporó al análisis de la comunicación, en diversas formas, el interés por los "estereotipos" sociales (LIPPMANN, 1922). Asimismo, adquirieron importancia cuestiones como el modo en que se presentaba a los negros en la prensa de Filadelfia (SIMPSON, 1934), el tratamiento otorgado en los textos de historia de los Estados Unidos a las guerras en las que participó este país, en comparación con el de sus ex enemigos (WALWORTH, 1938), o la expresión del nacionalismo en los libros infantiles estadounidenses, británicos y de otros países europeos (MARTIN, 1936).

Uno de los conceptos claves que apareció en la psicología de esta época fue el de "actitud". Este concepto añadió al análisis de contenido dimensiones evaluativas como las de los "adeptos" y los "contrarios" o las actitudes "favorables-desfavorables"; y esto, junto a los posteriores avances en el desarrollo de la teoría de las actitudes, abrió las puertas a la evaluación sistemática de las orientaciones, recurriendo a patrones como los de "objetividad", "equidad" y "equilibrio". Entre los criterios explícitos, vale la pena mencionar el "coeficiente de desequilibrio", de JANIS y FADNER (1965). Los experimentos psicológicos sobre transmisión de rumores llevaron a Allport y Fadner a analizar desde una perspectiva totalmente distinta el material periodístico; en su trabajo "Five Tentative Laws of the Psychology of Newspapers" (1940) intentaron dar cuenta de los cambios que experimenta la información a medida que se desplaza a través de una institución, hasta aparecer finalmente sobre la página impresa.

El interés por los símbolos políticos añadió otra característica al análisis de los mensajes públicos. MCDIARMID (1937), por ejemplo, analizó treinta discursos pronunciados por presidentes norteamericanos al inicio de su mandato en función de los símbolos de identidad nacional, de sus referencias históricas, de la alusión a conceptos fundamentales de la administración pública, y a hechos y expectativas. Y LASSWELL (1938), sobre todo, al examinar las comunicaciones públicas desde el punto de vista de su teoría psicoanalítica de la política, clasificó los símbolos en categorías como las correspondientes al "sí-mismo" a los "otros", a las formas de "indulgencia" y a la "carencia" o "privación". Su análisis de los símbolos dio origen a un "estudio mundial de la atención" (1941), en el cual se comparaban las tendencias registradas en la frecuencia de presentación de símbolos

nacionales, tal como aparecían en prestigiosos periódicos de diversos países.

Varias disciplinas examinaron sus respectivas tendencias académicas según se reflejaban en los temas aparecidos en los periódicos más representativos. Probablemente el primer examen de esta índole fue el relacionado con el desarrollo de la física en Rusia (RAINOFF, 1929), pero se llevó a cabo más cabalmente en el campo de la sociología (BECKER, 1910-1932; SHANAS, 1945) y luego se aplicó al periodismo (TANNENBAUM y GREENBERG, 1961).

Los rasgos que distinguen los primeros análisis de contenido del análisis cuantitativo de los periódicos son los siguientes: 1) muchos científicos sociales eminentes se incorporaron a esta evolución proporcionando ricos marcos teóricos; 2) se definieron, y se reconocieron en los datos correspondientes, conceptos bastante específicos, como los de actitud, estereotipo, estilo, símbolo, valor y métodos de propaganda; 3) se aplicaron al análisis herramientas estadísticas más perfectas, especialmente las procedentes de la investigación mediante encuestas y de los experimentos psicológicos; y 4) los datos provenientes del análisis de contenido pasaron a formar parte de trabajos de investigación de mayor envergadura (LAZARSFELD y otros, 1948). La primera exposición sumaria de estos trabajos apareció en el volumen titulado *Content Analysis in Communication Research*, del que originalmente fueron autores BERELSON y LAZARSFELD (1948), y más adelante únicamente BERELSON (1952).

ANALISIS DE LA PROPAGANDA

El análisis de contenido recibió un gran impulso gracias a la que probablemente fue su primera aplicación práctica de envergadura durante la segunda guerra mundial. Antes de la conflagración, el análisis de contenido se definía virtualmente por el uso de las comunicaciones de masas como datos para la corroboración de hipótesis científicas, campo en el cual su contribución fue importante, y por la crítica de las prácticas periodísticas. En este sentido, el análisis de la propaganda se inició como un instrumento para la identificación de los individuos que constituían fuentes de influencia "no éticas". De acuerdo con lo manifestado por el INSTITUTE FOR PROPAGANDA ANALYSIS (1937), los propagandistas se denuncian a sí mismos por el uso de ciertos procedimientos, como el empleo de

"improperios", el recurso a "generalidades llamativas", el llamar a las cosas y a las personas por su nombre, identificándolas sin rodeos, el hecho de "trucar los dados" en su favor o de adherirse siempre a la causa ganadora, etc. Estos procedimientos eran fácilmente discernibles en los discursos religiosos o políticos.

En la década de los 40, cuando en los Estados Unidos toda la atención estaba centrada en el esfuerzo bélico, la identificación del propagandista dejó de ser un problema. Tampoco bastó ya con afirmar que los medios de comunicación de masas eran agentes eficaces para manipular la opinión pública: era necesario llevar a cabo operaciones de inteligencia militar y política. En este clima surgieron dos destacados centros de análisis de la propaganda. Harold D. Lasswell y sus colaboradores comenzaron a trabajar en la Experimental Division for the Study of Wartime Communications de la Biblioteca del Congreso de los Estados Unidos, mientras que Hans Speier, que había organizado un proyecto de investigación sobre las comunicaciones totalitarias en la New School for Social Research, reunió un equipo de investigadores en el Foreign Broadcast Intelligence Service, perteneciente a la American Federal Communications Commission (FCC). La labor del grupo de la Biblioteca del Congreso (LASSWELL y otros, 1965) estuvo centrada en las cuestiones básicas del muestreo, los problemas de la medición, la fiabilidad y validez de las categorías de contenido establecidas y, en cierto sentido, prosiguieron con la tradición del primitivo análisis cuantitativo de las comunicaciones de masas. El grupo de la FCC, entretanto, basándose en las emisiones radiofónicas internas del enemigo, trató de comprender y predecir los sucesos que tenían lugar en el interior de la Alemania nazi y sus países aliados, así como de estimar los efectos que las acciones militares tenían sobre el espíritu bélico de sus poblaciones. La urgencia de la información cotidiana dejaba poco tiempo a los analistas para formalizar sus métodos, pero, tras la contienda, GEORGE (1959a) sometió al análisis los volúmenes de informes de guerra que habían sobrevivido, para verificar los métodos desarrollados en el curso de ese proceso y validar las inferencias efectuadas con los documentos de los que entonces se disponía. Estos trabajos dieron como resultado su *Propaganda Analysis*, una obra que supuso contribución fundamental para la conceptualización de los objetivos y procedimientos del análisis del contenido. La premisa de que el propagandista obra de manera racional, en el sentido de que se atiene a su propia teoría de la

propaganda al elegir sus aseveraciones, reorientó el eje del análisis, apartándolo del concepto neutral de "contenido compartido" para poner el acento en las condiciones que explican una comunicación y los intereses que esas condiciones favorecen presumiblemente. Una herramienta particularmente útil para comprender los significados de esas emisiones radiofónicas fue el concepto de "propaganda preparatoria": con el fin de asegurar el apoyo a una medida política prevista, la población afectada por ella debe ser preparada de manera que dicha medida le resulte aceptable. Los analistas de la FCC lograron predecir con éxito varias campañas militares y políticas importantes, evaluando las intuiciones de la cúspide nazi con respecto a su situación, los cambios políticos acaecidos en el interior del grupo dirigente y la modificación de las relaciones entre los países del Eje. Entre las predicciones más notables de los analistas británicos debe mencionarse la fecha del despliegue de los cohetes V alemanes de largo alcance contra Gran Bretaña. Mediante el estudio de los discursos de Goebbels, el analista pudo inferir las interferencias que habían existido en la producción de dichas armas y extrapolar con exactitud la fecha de su lanzamiento, semanas después.

Entre las lecciones que supuso esta aplicación se encuentran las siguientes:

1. El contenido no constituye una cualidad absoluta u objetiva de las comunicaciones. Emisor y receptor pueden diferir radicalmente en su manera de interpretar ciertas emisiones radiofónicas. Además, la secuencia de éstas y la situación particular en que se encuentra el receptor son determinantes para el significado de las comunicaciones. En otras palabras, el aspecto del contenido que comparten todos los comunicadores carece comparativamente de importancia para la comprensión del proceso político implicado.
2. El analista, que "mira por debajo de la superficie" de los mensajes propagandísticos, formula predicciones o inferencias acerca de los fenómenos sin tener un acceso directo a ellos. Ahora bien: el contenido no es identificable del mismo modo en que lo son las huellas dactilares. Los especialistas que habían sometido los periódicos a un análisis cuantitativo formularon inferencias, pero no las relacionaron con la situación de la que emergían sus datos, y erróneamente negaron su propia aportación a la comprensión de los mensajes que analizaban.
3. Con el fin de interpretar los mensajes propagandísticos o de dotarlos de sentido, es necesario contar con modelos elaborados sobre

el sistema en el cual tienen lugar las comunicaciones. El analista construía dichos modelos de manera más o menos explícita; pero mientras que en los primeros análisis del contenido las comunicaciones de masas se consideraban unidades aisladas, el análisis de la propaganda tiene en cuenta su naturaleza sistémica.

4. Los indicadores cuantitativos son muy poco sensibles y bastante burdos para suministrar interpretaciones políticas. Aunque se disponga de la gran cantidad de datos que exigen los análisis estadísticos, no conducen a las conclusiones "más obvias" que los expertos en política son capaces de extraer paralelamente observando con más profundidad las variaciones cualitativas.

Convencidos de que el análisis del contenido no debía ocupar un papel menor en la exploración del intelecto humano, numerosos autores (por ejemplo KRACAUER, 1947, 1952-1953; GEORGE, 1959) pusieron en tela de juicio la creencia simplista de que aquél debiera dedicarse a computar los datos cualitativos. SMYTHE (1954) sostuvo que esto revelaría la "inmadurez de la ciencia", al confundir objetividad con cuantificación. No obstante, los propugnadores del enfoque cuantitativo hicieron en gran medida caso omiso de esta crítica, y en su ensayo de 1949, "Why Be Quantitative", LASSWELL (1965a) continuó sosteniendo que la cuantificación de símbolos era la única base para obtener interpretaciones científicas.

GENERALIZACION DEL ANALISIS DE CONTENIDO

Después de la segunda guerra mundial, y quizá como consecuencia del primer panorama integral del análisis del contenido que suministraron BERELSON y LAZARSFELD (1948) y BERELSON (1952), el análisis del contenido se amplió a numerosas disciplinas. Si bien es cierto que la comunicación de masas dejó de ser su dominio empírico exclusivo, las aplicaciones referidas a ellas fueron, y siguen siendo, predominantes. De hecho, algunos de los mayores proyectos de investigación estaban relacionados con los medios de comunicación públicos. Por ejemplo, LASSWELL (1941) concretó su idea de un "estudio mundial sobre la atención" sometiendo al análisis los símbolos políticos de los editoriales de medios de prensa oficialistas de Francia, Alemania, Gran Bretaña, Rusia y los Estados Unidos, así como los que apa-

recían en discursos políticos importantes. Su intención era poner a prueba la hipótesis de que se había puesto en marcha una "revolución mundial" (LASSWELL y otros, 1952). La propuesta de GERBNER (1969) acerca de utilizar "indicadores culturales" fue ejemplificada en su análisis, llevado a cabo por décimo año consecutivo, de los programas televisivos de ficción emitidos durante una semana, principalmente con el objeto de establecer las dosis de violencia de distintas redes de televisión, investigar sus tendencias y ver de qué manera diversos grupos de la población (mujeres, niños, personas de edad avanzada, etc.) eran descritos en la televisión norteamericana (GERBNER y otros, 1979).

En el campo de la psicología, el análisis del contenido se centró en tres aplicaciones primordiales. La primera fue el análisis de los registros verbales para descubrir características motivacionales, psicológicas o de la personalidad. Esta aplicación instauró una tradición propia desde que ALLPORT (1942) publicó su tratado sobre el uso de los documentos personales, desde la publicación por parte de BALDWIN (1942) de su *Personality Structure Analysis* para verificar la estructura cognitiva, y desde los estudios de WHITE (1947) sobre los valores. Una segunda aplicación apareció determinada por la reunión de datos cualitativos en forma de respuestas a preguntas abiertas, respuestas verbales a los tests y elaboración de los relatos del Test de Aptitud Temática. En este caso, el análisis del contenido adquirió el carácter de una técnica complementaria, que permitía al investigador utilizar datos que sólo podían reunirse sin imponer al sujeto una estructuración demasiado excesiva, y convalidar entre sí los hallazgos obtenidos mediante distintas técnicas. La tercera aplicación estaba relacionada con los procesos de comunicación de los que el contenido forma parte integral. Por ejemplo, en su obra *Interaction Process Analysis*, BALES (1950) demostró que los intercambios verbales en pequeños grupos podían utilizarse como datos para analizar los procesos que se daban en ellos. Los antropólogos comenzaron a utilizar las técnicas del análisis de contenido para el examen de los mitos, leyendas y acertijos; el análisis componencial de la terminología del parentesco fue uno de los numerosos ejemplos de las contribuciones específicas de esta disciplina al análisis del contenido. Los historiadores siempre habían buscado métodos más sistemáticos para el examen de grandes conjuntos de documentos históricos disponibles, y vieron en el análisis de contenido el instrumento adecuado. El material educativo, que durante mucho tiempo fue el centro de

atención de los científicos sociales, pasó a constituir una rica fuente de datos tanto para realizar inferencias sobre los procesos de la lectura (FLESH, 1948-1951) como para comprender las tendencias más amplias, de carácter político, actitudinal y valorativo, que pueden aparecer en los libros de texto. El problema de la identificación de los autores de documentos no firmados (un problema de origen antiguo entre los eruditos de la literatura) fue abordado con éxito mediante estas nuevas técnicas; y así sucesivamente. Pero esta proliferación trajo consigo una pérdida de claridad: todo parecía susceptible de ser sometido al análisis del contenido, y todo análisis de fenómenos simbólicos se convirtió en un análisis de contenido.

En 1955, el Social Science Research Council's Committee on Linguistics and Psychology organizó una conferencia sobre el análisis del contenido, respondiendo de esta manera al creciente interés por este tema. Los participantes provenían de disciplinas como la psicología, las ciencias políticas, la literatura, la historia, la antropología y la lingüística. Sus aportaciones aparecieron en un volumen compilado por Pool, *Trends in Content Analysis* (1959). Pese a las obvias divergencias en cuanto al interés y enfoque de los participantes, Pool observaba una coincidencia considerable, y a menudo sorprendente, en lo que se refiere a dos aspectos: existía una aguda preocupación por los problemas de las inferencias realizadas a partir del material verbal con respecto a sus circunstancias antecedentes, y se insistía en el cómputo de las relaciones internas entre símbolos, y no en el cómputo de las frecuencias de aparición de los mismos símbolos (POOL, 1959, pág. 2).

ANALISIS DE TEXTOS POR ORDENADOR

A fines de la década de los 50 se produjo un considerable auge del interés por la traducción automática, la preparación automática de resúmenes y los sistemas mecánicos de recuperación de información. Se desarrollaron lenguajes de computación especialmente apropiados para el procesamiento de datos literales, e incluso aparecieron revistas especializadas en las aplicaciones del ordenador a la psicología, las humanidades y las ciencias sociales. El volumen a menudo cuantioso de documentos escritos que debían analizarse, así como la repetitividad de esta tarea, convertían al ordenador en un aliado natural de los especialistas en aná-

lisis del contenido.

Una vez que la evolución de los soportes lógicos *[software]* convirtió al ordenador en un instrumento cada vez más eficaz para el procesamiento de datos alfabéticos (en oposición a los numéricos), se pudo contar con programas para el cómputo de palabras, los cuales sirvieron de base, entre otras cosas, para una nueva disciplina, posteriormente llamada "estilística computacional" (SEDELOW y SEDELOW, 1966) y produjeron una revolución en las tediosas tareas relacionadas con las obras literarias (por ejemplo, el establecimiento de concordancias). Probablemente fueron SEBEOK y ZEPS (1958) quienes abordaron primero un análisis de contenido por ordenador en el que se aplicaron rutinas de recuperación de información para analizar unas cuatro mil leyendas populares de los cheremis.* El trabajo de HAYS (1960) *Automatic Content Analysis*, publicado por la Rand Corporation, exploraba la posibilidad de crear un sistema de computación para el análisis de documentos políticos. Ignorantes de estos avances, Stone y Bales, que estaban llevando a cabo un estudio temático de grupos de interacción cara a cara, diseñaron y programaron la versión inicial del sistema General Inquirer. Una versión posterior de este sistema, así como numerosas aplicaciones que iban desde la ciencia política hasta la publicidad y desde la psicoterapia hasta el análisis literario, fueron presentadas por STONE y otros (1966). Desde esa fecha, el sistema ha sido ampliado y perfeccionado y se ha hecho extensivo a otros ámbitos lingüísticos distintos del inglés.

Los usos del ordenador en el análisis de contenido se vieron estimulados también por los desarrollos que tuvieron lugar en otros campos. La psicología se interesó por simular la cognición humana (ABELSON, 1963). NEWELL y SIMON (1963) crearon un método computacional para la resolución de problemas (humanos). En la lingüística surgieron numerosos enfoques de análisis sintáctico y la interpretación semántica de las expresiones lingüísticas. La inteligencia artificial se centró en la creación de máquinas que pudieran comprender los lenguajes naturales.

En 1967, la Escuela de Comunicaciones de Annenberg auspició una importante conferencia sobre el análisis de contenido, en la que investigadores pertenecientes a numerosas disciplinas tuvieron la oportunidad de presentar sus propias técnicas compu-

* Los cheremis son un pueblo finés oriental de la región de Kazán, en la república soviética de Tartaria. [E.]

tacionales, comparar la eficacia de distintos enfoques y proyectar las necesidades futuras. Además, la conferencia sirvió como plataforma para cotejar los diversos enfoques sobre el análisis de contenido. Los debates giraron en torno a las dificultades para registrar las comunicaciones no verbales (visuales, vocales y musicales), la necesidad de establecer categorías estandarizadas, el problema de la obtención de inferencias y, en particular, el papel de las teorías y de las construcciones analíticas, todo lo cual plantea, en esencia, problemas metodológicos y no computacionales. BARCUS (1959), que había reunido gran parte de la bibliografía sobre el análisis del contenido, expuso una reseña acerca de su uso en la investigación y la docencia en los Estados Unidos. Estas contribuciones fueron sintetizadas en un volumen compilado por GERBNER y otros (1969), cuya publicación coincidió con una revisión del estado de la cuestión sobre este campo realizada por HOLSTI (1969).

En síntesis, se puede afirmar que el análisis del contenido ha llegado a ser un método científico capaz de ofrecer inferencias a partir de datos esencialmente verbales, simbólicos o comunicativos. Más allá de su continuo compromiso con cuestiones psicológicas, sociológicas y políticas sustanciales, en los últimos ochenta años ha aumentado de forma exponencial el interés por el uso de esta técnica y se ha procurado establecer criterios adecuados de validez. Consideramos que esto indica una madurez cada vez mayor.

2. Fundamentos conceptuales

El análisis de contenido tiene su propio método para analizar los datos, que procede en gran medida de su manera de considerar el objeto de análisis, es decir, el contenido. En este capítulo se exponen los fundamentos conceptuales del análisis de contenido, se ofrece una definición de este último, se desarrolla un marco de referencia con fines prescriptivos, analíticos y metodológicos, y se compara el análisis de contenido con otras técnicas de la ciencia social.

Definicion

Proporcionaremos la siguiente definición:

El análisis de contenido es una técnica de investigación destinada a formular, a partir de ciertos datos, inferencias reproducibles y válidas que puedan aplicarse a su contexto.

Como técnica de investigación, el análisis de contenido comprende procedimientos especiales para el procesamiento de datos científicos. Al igual que todas las restantes técnicas de investigación, su finalidad consiste en proporcionar conocimientos, nuevas intelecciones, una representación de los "hechos" y una guía práctica para la acción. Es una herramienta.

De cualquier instrumento de la ciencia se espera que sea fia-

ble. Más concretamente, si otros investigadores, en distintos momentos y quizás en diferentes circunstancias, aplican la misma técnica a los mismos datos, sus resultados deben ser los mismos que se obtuvieron originalmente. Este es el requisito que se tiene en cuenta al decir que el análisis de contenido debe ser *reproducible*.

Otra definición, la de BERELSON (1952), sostiene que el análisis de contenido es "una técnica de investigación para la descripción objetiva, sistemática y cuantitativa del contenido manifiesto de la comunicación" (pág, 18). Por supuesto, el requisito de que la técnica sea "objetiva" y "sistemática" está incluido en nuestra definición, cuando hablamos del requisito de reproducibilidad. Para que un proceso sea reproducible, las reglas que lo gobiernan deben ser explícitas e igualmente aplicables a todas las unidades de análisis. No obstante, hemos excluido de nuestra definición varios requisitos que Berelson plantea en la suya, en gran medida porque son, o bien poco claros, o demasiado restrictivos. Por ejemplo, Berelson incorpora el atributo de que el contenido sea "manifiesto" simplemente para asegurar que la codificación de los datos en el análisis de contenido sea intersubjetivamente verificable y fiable. Esta definición ha provocado que muchos estudiosos piensen que los contenidos latentes están excluidos del análisis. La exigencia de que la descripción sea "cuantitativa" es análogamente restrictiva: si bien en muchas actividades científicas la cuantificación es importante, los métodos cualitativos han demostrado su eficacia, particularmente en lo que se refiere a extraer de la propaganda información utilizable con fines políticos o militares, así como en psicoterapia y, aunque parezca extraño, en el análisis por ordenador de datos lingüísticos, donde las consideraciones cualitativas resultaron fundamentales para el establecimiento de algoritmos adecuados.

La principal objeción que puede hacerse a la definición de Berelson es que no explica en qué consiste el "contenido" o cuál debería ser el objeto de un análisis de contenido. Para algunos investigadores, el "análisis de contenido" no parece designar otra cosa que el cómputo de cualidades (palabras, atributos, colores): para otros, esa expresión sugiere la existencia de un método para "extraer" contenidos de los datos, como si estuviesen objetivamente "contenidos" en éstos. Ninguna de estas concepciones aborda el núcleo del problema del análisis de contenido.

Nuestra definición procura ser explícita en lo relativo al objeto

del análisis de contenido; pero dado que esto tal vez no resulte inmediatamente claro, procederemos por etapas.

Intuitivamente, el análisis de contenido podría caracterizarse como un método de investigación del *significado simbólico* de los mensajes. La mayoría de los expertos en análisis de contenido probablemente tengan esta idea en mente, y dicha caracterización podría parecer razonable si no fuese por dos connotaciones equívocas, por lo menos, que una buena definición debe evitar.

En primer lugar, *los mensajes no tienen un único significado* que necesite "desplegarse". Siempre será posible contemplar los datos desde múltiples perspectivas, en especial si son de naturaleza simbólica. En cualquier mensaje escrito se pueden computar letras, palabras u oraciones; pueden categorizarse las frases, describir la estructura lógica de las expresiones, verificar las asociaciones, denotaciones, connotaciones o fuerzas ilocutivas; y también pueden formularse interpretaciones psiquiátricas, sociológicas o políticas. Todas estas cosas pueden poseer validez de forma simultánea. En suma, un mensaje es capaz de transmitir una multiplicidad de contenidos incluso a un único receptor. En estas circunstancias, la pretensión de haber analizado *el* contenido de la comunicación trasluce una posición insostenible.

En segundo lugar, *no es necesario que exista coincidencia acerca de los significados.* Si bien el consenso o el acuerdo intersubjetivo sobre lo que significa un mensaje simplifica enormemente el análisis de su contenido, dicho acuerdo sólo existe en relación con los aspectos más obvios o "manifiestos" de las comunicaciones, o bien para unas pocas personas que comparten la misma perspectiva cultural y sociopolítica. En general, ninguna de estas condiciones reviste interés. Así, pues, difícilmente la coincidencia puede servir como presupuesto para un análisis de contenido. En las interacciones entre un psiquiatra y su paciente, un especialista avezado conversa con un lego acerca de los problemas de este último: no puede presumirse que sus perspectivas sean iguales. Incluso los expertos en artefactos antropológicos, o en arte, comunicación no verbal y política, suelen discrepar con sus informantes (los nativos o los participantes en algunas de esas actividades) acerca de la forma en que deben interpretarse los símbolos que ellos utilizan. Los oradores, o aquellos que pronuncian discursos públicos tienden a emplear expresiones ambiguas de forma premeditada, poniendo así de manifiesto su conciencia asimétrica del hecho de que los mensajes son capaces

de transmitir distintas cosas a distintas personas. En consecuencia, los significados se refieren siempre a un comunicador.

El rasgo más característico de los mensajes es que informan a una persona de manera vicaria, proporcionando al receptor un conocimiento acerca de hechos o sucesos que se producen en un lugar lejano, acerca de objetos que quizás existieron en el pasado o acerca de las ideas de otras personas. *Los mensajes y las comunicaciones simbólicas tratan, en general, de fenómenos distintos de aquellos que son directamente observados.* La naturaleza vicaria de las comunicaciones simbólicas es lo que obliga al receptor a formular inferencias específicas, a partir de los datos que le proporcionan sus sentidos, en relación con ciertas porciones de su medio empírico. A este medio empírico lo denominamos *el contexto de los datos.* Asimismo, advertimos que siempre es *una persona concreta* quien formula esas inferencias a partir de los datos sobre su contexto; y, por ello, es esa persona quien distingue si sus experiencias son vicarias o directas, si algo es simbólico o no simbólico, o si el dato de que dispone es un mensaje sobre alguna otra cosa o es un hecho que despliega su propia existencia y estructura. El analista del contenido es también un receptor de datos, aunque probablemente difiera de manera radical de los comunicadores comunes, que quizás asignen significados de forma rutinaria e inconsciente, y sin justificación empírica. Si bien el análisis del contenido puede ocuparse de *formular la clase de inferencias que efectúa algún receptor cuando trata de comprender las comunicaciones simbólicas*, la técnica ha sido generalizada, y alcanza probablemente su mayor grado de éxito al aplicarla a formas no lingüísticas de comunicación, donde las pautas presentes en los datos son interpretadas como *índices y síntomas*, de los cuales los comunicadores no avezados tal vez ya no sean conscientes.

La formulación de inferencias específicas —que es la clave que distingue entre el procesamiento de datos simbólicos y de datos no simbólicos, delineando así el ámbito del análisis del contenido— no está enteramente ausente de otras definiciones. Por ejemplo, Holsti y Stone abogan por una definición que se aparta de la de Berelson por lo menos en dos aspectos importantes. En primer lugar, reconoce el carácter inferencial de la codificación de unidades textuales en categorías conceptuales. En segundo lugar, convierte las inferencias en su principal preocupación: "El análisis del contenido es una técnica de investigación

para formular inferencias identificando de manera sistemática y objetiva ciertas características especificadas dentro de un texto" (STONE y otros, 1966, pág. 5). Si bien esta definición reconoce la naturaleza inferencial de la identificación de las formas de las ideas, valores y actitudes a los que se aplica el análisis del contenido, no explica la importancia de relacionar dicha clasificación, categorización y cómputo de la frecuencia de estas formas con otros fenómenos. Y esto es necesario si se quiere que los resultados del análisis de contenido tengan un significado empírico. Ya en 1943, JANIS (1965) apuntaba la necesidad de convalidar los resultados del análisis de contenido de la comunicación de masas relacionándolos con las percepciones del público o con los efectos en el comportamiento de éste. También nosotros exigimos que el análisis de contenido sea predictivo de algo en principio observable, que facilite la toma de decisiones o que contribuya a conceptualizar la porción de la realidad que dio origen al texto analizado. Con este fin, sugerimos que cualquier análisis de contenido debe realizarse en relación con *el contexto de los datos*, y justificarse en función de éste.

Creemos, además, que todas las teorías del significado, todas las teorías de los fenómenos simbólicos, incluidas las que se ocupan del contenido del mensaje, se asemejan entre sí por la importancia que conceden a la relación entre los datos y su contexto. En su sentido más elemental, los datos son estímulos físicos o vehículos-signos, como las marcas negras sobre un papel blanco. Sin embargo, el interés por el significado parte casi siempre de niveles superiores de abstracción, como documentos escritos, películas cinematográficas, diálogos verbales y pinturas, por mencionar sólo unos pocos casos. El contexto es su medio. Habitualmente, el analista está en condiciones de elegir el medio y la manera de conceptualizarlo. Un lingüista puede limitar su foco de atención al entorno lingüístico de las palabras y expresiones. Un sociólogo quizá reconozca el significado de un acto situándolo dentro del contexto social de la situación en la que ocurrió. Un investigador de las comunicaciones tal vez interprete el significado de un mensaje en relación con las intenciones del emisor, los efectos cognitivos o conductuales que ejerce sobre el receptor, las instituciones dentro de las cuales se produjo el intercambio o la cultura en la que desempeña un papel.

ELABORACIONES

Sin entrar en un posterior análisis de los términos en que se expresan los mensajes, algunos ejemplos podrán aclarar mejor de qué manera la definición se aplica a situaciones prácticas. Considérese el caso del analista de emisiones radiofónicas del enemigo en época de guerra, que procura evaluar el grado de apoyo popular que tienen las políticas adoptadas por los círculos gobernantes del país enemigo. En época de paz, un problema de esta índole quizá podría resolverse mediante una encuesta de opinión pública o una inspección *in situ.* Durante la guerra, dicha información, por lo general, es inaccesible, y el analista se ve obligado a obtenerla por medios indirectos. En este caso al analista no le interesa lo que el propagandista intenta transmitir, ni tampoco le molesta que sus inferencias no sean compartidas por los oyentes ordinarios. En realidad, puede que tenga buenos motivos para descartar los significados manifiestos de los mensajes, considerando que éstos mienten. Lo que no puede ignorar es que las "emisiones radiofónicas del enemigo" forman parte de un proceso político inaccesible pero real, que incluye una población civil, su círculo gobernante, los militares y las condiciones sociales, políticas y económicas del país. Por propia decisión del analista, los actores de este proceso constituyen el contexto de las emisiones radiofónicas. Dentro de ese contexto se sitúan los fenómenos que le interesan, a saber: el grado de apoyo popular que reciben las políticas de los gobernantes. Con el objeto de formular una inferencia válida acerca de dicho apoyo popular a partir de las emisiones radiofónicas que se difunden en el país, el analista debe conocer en alguna medida las relaciones existentes entre los principales elementos que intervienen en el proceso político.

Tampoco los historiadores son jamás simples recopiladores de documentos. La fascinación que producen procede de la reconstrucción simbólica de sucesos del pasado, a la que es posible acceder gracias a los textos disponibles. Sin duda, todo historiador debe poseer el profundo anhelo de entrevistar a César, a Nietzsche o a los esclavos africanos que entraron en América durante la época colonial, si ello fuera posible; pero dado que las cartas, libros, instrumentos y otros registros elaborados por otras personas, no prevén cuáles han de ser los interrogantes del historiador, éste tiene que buscar respuestas por métodos indirectos. Cuando un historiador se empeña en inferir acontecimientos a

partir de documentos (DIBBLE, 1963), está practicando, por definición, un análisis de contenido. Por lo tanto, no ha de sorprender que los historiadores siempre hayan exigido que los documentos se sitúen dentro del contexto histórico apropiado. Una vez delineado este contexto, las brechas en los detalles pertinentes se llenan extrayendo inferencias a partir de numerosos fragmentos de información. Los métodos históricos contribuyen a crear una red de relaciones que, en última instancia, pueda dar respuesta a los interrogantes originalmente planteados. El proceso de disminución de la incertidumbre en un ámbito no observable es de carácter inferencial, y de hecho es el mismo que sigue el análisis de contenido.

La psicología tiene una larga tradición en lo referente al desarrollo de teorías y la construcción de pruebas empíricas con un grado de generalidad establecido mediante experimentos reiterados. El análisis de documentos personales que llevan a cabo los psicólogos (ALLPORT, 1942) se ajusta claramente a la descripción del análisis de contenido: su finalidad y su contexto los lleva a ocuparse de temas como la cognición del autor del documento, sus actitudes, personalidad o psicopatología, todo lo cual podría haberse medido mediante técnicas más directas si el autor estuviera disponible. En el curso del análisis de los documentos personales han surgido una variedad de técnicas de inferencia (por ejemplo, la razón tipo figuración, el cociente de alivio de la incomodidad, la grafología, la norma de legibilidad, los tests de apercepción temática) destinadas a relacionar la sintaxis y las expresiones que aparecen en los documentos escritos con características propias de su autor. En psicología, el análisis del contenido es una práctica cotidiana, aunque tal vez no siempre se haya etiquetado de esa manera.

Por supuesto, el dominio tradicional del análisis de contenido ha sido el de la comunicación de masas. Los investigadores suelen interesarse por averiguar las características del comunicador, los efectos que su mensaje ejerce sobre el auditorio, el grado de tensión pública, el clima sociopolítico, los procesos de mediación de los valores, los prejuicios, las diferenciaciones culturales, las limitaciones institucionales, etc. Cuando un análisis de contenido se ocupa, por ejemplo, de las percepciones del público, está claramente implícito el contexto de la comunicación de masas; pero éste resulta menos claro cuando el investigador sólo pretende describir en qué consisten las comunicaciones. Entonces parece

que toda inferencia ha sido omitida, y que la finalidad descriptiva no puede satisfacer nuestra definición del análisis del contenido. Pero no es así.

Veamos el caso de las elucidaciones "puramente descriptivas", según se las llama, de las tendencias políticas, el prejuicio racial o la violencia televisiva. Si bien dichas elucidaciones pueden presentarse como "fácticas", sólo tienen un significado dentro del contexto de los problemas sociales que les otorgan pertinencia. Sería muy raro que alguien emprendiera una descripción sistemática de algo sin tener en mente alguna intención o implicación. La forma en que un analista conceptualiza y describe las tendencias u orientaciones, los prejuicios y la violencia, nunca es del todo arbitraria. Tal vez el lenguaje descriptivo utilizado sea el de la institución o persona que le encomendó el análisis de contenido; tal vez se acomode a las percepciones de un público determinado o a los intereses de un productor de programas para una cadena radiofónica o televisiva, o quizá refleje los estereotipos culturales o las convenciones literarias vigentes. Sea como fuere, *toda descripción supone inferencias*, por primitiva que sea. La cuantificación de la violencia televisiva, por ejemplo, difícilmente podría justificarse si no pusiera de manifiesto una correlación con otros fenómenos (conducta agresiva o delictiva), o si no pretendiera mostrar alguna correspondencia con las percepciones de los integrantes de un público determinado, u ofrecer algún apoyo firme para favorecer la aplicación de una política. El análisis de contenido, aun cuando tenga finalidades sólo descriptivas, no debe ser ajeno a las consideraciones relacionadas con la validez, y tiene que estar específicamente vinculado con el contexto al que pertenecen los hallazgos.

Marco de referencia conceptual

La definición del análisis de contenido establece el objeto de investigación y sitúa al investigador en una posición concreta frente a su realidad. Basándonos en trabajos anteriores (Krippendorff, 1969a, págs. 7-13), ofrecemos a continuación un marco de referencia conceptual dentro del cual puede representarse el papel que desempeña el investigador. Este marco es simple y general, y en él sólo se recurre a unos pocos conceptos básicos:

- Los *datos*, tal como se comunican al analista.
- El *contexto* de los datos.
- La forma en que *el conocimiento del analista* lo obliga a dividir su realidad.
- El *objetivo* de un análisis de contenido.
- La *inferencia* como tarea intelectual básica.
- La *validez* como criterio supremo de éxito.

Este marco de referencia tiene tres finalidades: es prescriptivo, analítico y metodológico. Es prescriptivo en el sentido de que debe guiar *la conceptualización y el diseño* de los análisis de contenido prácticos en cualquier circunstancia; es analítico en el sentido de que debe facilitar *el examen crítico* de los resultados del análisis de contenido efectuado por otros; y es metodológico en el sentido de que debe orientar *el desarrollo y perfeccionamiento sistemático de los métodos* de análisis de contenido.

En todo análisis de contenido *debe quedar claro qué datos se analizan*, de qué manera se definen y de qué población se extraen. Los datos son lo único disponible para el especialista en análisis de contenidos, y no su contexto. Los datos exhiben su propia sintaxis y estructura, y se describen en función de unidades, categorías y variables, o son codificados de acuerdo con un esquema multidimensional. Los datos son los elementos básicos, primitivos, del análisis de contenido, y constituyen la superficie que el analista debe tratar de penetrar. *La comunicación de los datos al analista es unidireccional.* El es incapaz de manipular la realidad; no dispone de una realimentación correctiva con la fuente que, por razones propias, le suministra información; se ve obligado, pues, a estudiar una porción de su universo de manera discreta. Por ejemplo, el analista de la propaganda bélica no puede interaccionar con los miembros de los grupos dirigentes del enemigo que le interesan; tampoco el análisis de contenido de los medios de comunicación puede influir en la situación en la cual se producen, difunden y reciben nuevas noticias, nuevos espectáculos, que a su vez se convierten en una variedad de conducta del público (ni tampoco puede recrear dicha situación).

En todo análisis de contenido *debe hacerse explícito el contexto con respecto al cual se analizan los datos.* Si bien los datos aparecen disponibles de una manera directa, su contexto lo construye el analista con el fin de incluir todas las condiciones circundantes, antecedentes, coexistentes o consecuentes. La necesidad

de delinear el contexto del análisis es particularmente importante porque no existen límites lógicos en cuanto al tipo de contexto que un analista puede querer considerar. Cualquier trabajo de investigación debe definir los límites más allá de los cuales no podrá extenderse el análisis. El análisis de contenido, las convenciones de cada disciplina y los problemas prácticos dictaminan con frecuencia la elección de esos límites. Los psicólogos sólo suelen interesarse en un autor individual, un político, una personalidad histórica o un paciente psiquiátrico. Los investigadores de la comunicación suelen situar los mensajes en el contexto de la interacción entre emisor y receptor. El analista de emisiones radiofónicas enemigas tal vez ya tenga en mente ciertos aspectos de la nación enemiga, en cuyo contexto cobrará sentido la conducta propagandística de su elite gobernante. Si bien los límites de un contexto son arbitrarios, es beneficioso definirlos de modo que exista alguna unidad estructural, alguna manera natural de dividir el universo entre lo que es pertinente y lo que no lo es.

En cualquier análisis de contenido, los intereses y conocimientos del analista determinan la construcción del contexto dentro del cual realizará sus inferencias. Por lo tanto, es importante que conozca el origen de sus datos y ponga de manifiesto los supuestos que formula acerca de ellos y acerca de su interacción con el medio. Debe ser capaz de distinguir entre dos clases de conocimiento: en primer lugar, aquel cuya naturaleza es variable o inestable, y con respecto a cuyo estado, forma o valor el analista no puede tener, en consecuencia, certidumbre alguna, y en segundo lugar, debe saber que existen ciertas relaciones entre las variables que son permanentes, fijas o estables. Como sucede con todo conocimiento, esta distinción se modificará en el transcurso del tiempo. En realidad, los análisis de contenido aplican los datos y conocimientos disponibles sobre las configuraciones estables con el fin de disipar la incertidumbre acerca de la pauta inestable que presenta el contexto de sus datos.

En todo análisis de contenido, *debe enunciarse con claridad la finalidad u objetivo de las inferencias.* El objetivo es lo que el analista quiere conocer. Dado que el análisis de contenido suministra un conocimiento vicario, información acerca de algo que no puede observarse directamente, ese objetivo se sitúa dentro de la porción variable del contexto de los datos disponibles. Es cierto que pueden realizarse muchos estudios exploratorios, en cuyo transcurso el investigador decidirá en qué centrará su atención;

pero al final tiene que adoptar un rumbo claro. Sólo si el objetivo del análisis de contenido es enunciado de manera inequívoca podrá juzgar si lo ha podido completar y aclarar el tipo de prueba que finalmente sea necesaria para convalidar los resultados.

La pretensión de haber analizado una muestra representativa del *New York Times* puede ser específica respecto de los datos, pero ambigua en lo referente al contexto y al objetivo perseguido. De hecho, puede que el análisis haya olvidado por completo la característica simbólica de este medio. Un título como "Análisis de contenido de los carteles en las paredes en China" es análogamente incierto, mientras que "Predicción de las decisiones del círculo gobernante a partir de los carteles en las paredes en China", o "Lo que aprende el público chino sobre la adopción de decisiones políticas a partir de los carteles de las paredes", hacen al menos referencia a algún contexto e indican una finalidad.

En todo análisis de contenido la tarea consiste en *formular inferencias*, a partir de los datos, en relación con algunos aspectos de su contexto, y justificar esas inferencias en función de lo que se sabe acerca de los factores estables del sistema en cuestión. Mediante este proceso se reconocen los datos como simbólicos o como susceptibles de proporcionar información acerca de algo que le interesa al analista.

Para llevar a cabo o justificar esas inferencias, el analista del contenido debe contar con relaciones relativamente estables entre los datos y el contexto (o construir una teoría operacional de esas relaciones), incluida la aportación de los factores mediadores. Se llama *construcción analítica* a una teoría de ésas relaciones, formuladas de tal manera que los datos aparezcan como sus variables independientes, y el objetivo forme parte de sus variables dependientes. Así, pues, una construcción analítica ofrece reglas de inferencia, sirviendo como puente lógico entre los datos disponibles y la finalidad incierta situada en su contexto. No basta con saber que las relaciones entre los datos y el contexto son estables; es preciso conocer también su forma específica. Al tratar de predecir el desenlace de una decisión política partiendo de las noticias que ofrecen los medios de comunicación de masas, no basta con saber que éstos desempeñan un papel importante en el proceso político: es preciso conocer qué elementos, aseveraciones o argumentos influyen en dicho proceso en una dirección o en otra. Lo que se necesita es conocer la naturaleza de esta relación.

En todo análisis de contenido *hay que especificar por adelantado el tipo de pruebas necesarias para validar sus resultados*, o hacerlo con la suficiente claridad como para que la validación resulte concebible. Si bien la razón de ser del análisis de contenido es justamente la falta de pruebas directas sobre los fenómenos en juego, los cuales deben por ello ser objeto de inferencia, al menos debe contarse con criterios claros para una validación de los resultados, con el fin de que otros individuos puedan recoger las pruebas adecuadas y comprobar si las inferencias son de hecho exactas. Esta validación *ex post facto* no es una simple cuestión de curiosidad. De la misma manera que sólo se puede aprender cuando se distingue entre el éxito y el fracaso, así también la metodología del análisis de contenido sólo puede progresar si se efectúan esfuerzos sistemáticos para validar sus resultados. Con mucha frecuencia los análisis de contenido se consideran casos singulares, y jamás se intenta su repetición, o bien se conceptualizan de manera tan insuficiente que resulta imposible pedirle al investigador que aclare cómo sería una prueba adecuada de validez. Estos análisis de contenido son empíricamente inútiles y contribuyen muy poco a la metodología del análisis del contenido.

Un buen ejemplo de una validación *ex post facto* se halla en el trabajo de GEORGE (1959a), al examinar los documentos incautados tras la segunda guerra mundial, con el objeto de averiguar hasta qué punto habían sido acertadas las aseveraciones de los analistas de propaganda de la FCC, y evaluar así las diferentes técnicas que utilizaron. JANIS (1965) propuso un método indirecto para validar las percepciones inferidas del público, correlacionando los resultados del análisis de contenido con el comportamiento de este último (por ejemplo, su conducta electoral, su consumismo o su agresividad).

Este marco de referencia se sintetiza en la figura 1. En ella se sugiere que los datos se disocian de sus fuentes o de sus condiciones circundantes y se comunican al analista de manera unidireccional. Este los sitúa en un contexto que construye basándose en su conocimiento de esas condiciones circundantes de los datos, incluido lo que desea conocer acerca del objetivo del análisis. Sus conocimientos sobre las relaciones estables dentro del sistema de interés se formulan como construcciones analíticas que le permiten extraer inferencias concernientes al contexto de los datos. Los resultados del análisis de contenido deben repre-

sentar alguna característica de la realidad, y la naturaleza de esa representación debe ser en principio verificable.

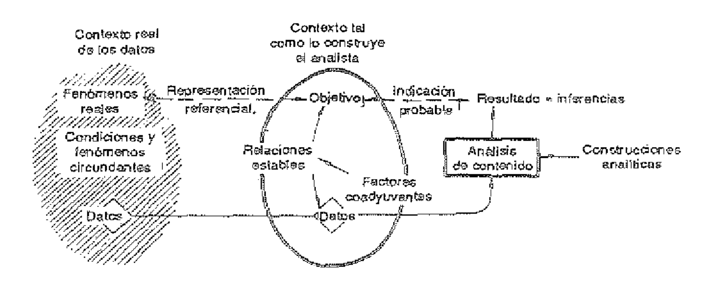


Figura 1. Marco de referencia para el análisis de contenido.

DISTINCIONES

Toda técnica de investigación tiene su propio ámbito empírico. Puede utilizarse incorrectamente, aplicándola allí donde otros métodos quizá resultarían más eficientes. Con el objeto de que el análisis de contenido se aproveche al máximo de sus posibilidades potenciales, compararemos su ámbito empírico con el de otras técnicas de investigación social. En la figura 2 se sintetizan estas distinciones, recurriendo a las tres dimensiones más importantes de contraste. Allí aparece el análisis de contenido en uno de los vértices del cubo, en tanto que los restantes métodos se hallan dispersos en distintos puntos. Enunciaremos estos contrastes en forma de proposiciones.

1. *El análisis de contenido no es una técnica intromisiva.* Desde Heisenberg sabemos que los actos de medición que intervienen en el comportamiento de los fenómenos que se desean estudiar crean observaciones contaminadas, tanto más cuanto más profundo sea el sondeo del observador. En lo que se refiere a las ciencias sociales. WEBB y otros (1966) han enumerado las diversas maneras en que los sujetos reaccionan cuando se les somete a observación científica, introduciendo así errores en los datos que se analizan; algunas de esas maneras son:

- La conciencia de ser observado o sometido a prueba.
- El rol que el sujeto asume o que le es asignado como entrevistado o informante.
- Las influencias que el proceso de medición ejerce sobre el sujeto.
- Los estereotipos y preferencias en la formulación de las respuestas.
- Los efectos de la interacción del experimentador-entrevistador sobre el sujeto.

Técnicas particularmente sensibles a estos errores son los experimentos, las entrevistas, los cuestionarios y los tests proyectivos. Todas ellas son conducidas por un investigador que controla en distintos grados las condiciones de estímulo frente a las cuales deben reaccionar los sujetos.

En contraste con ello, el análisis de contenido, la recuperación de información, la modelización, el uso de registros estadísticos y, hasta cierto punto, la etnometodología, tienen en común el hecho de constituir técnicas de investigación no reactivas o no intromisivas.

Hay dos razones por las cuales un investigador puede tratar de eludir las situaciones reactivas. *Puede pensar que la imposición de limitaciones excesivas a la situación que da origen a los datos puede poner en peligro su validez.* Por este motivo, los etnometodólogos prefieren registrar datos partiendo de medios naturales, los psiquiatras evitan formular preguntas que distraigan la atención del paciente de lo que éste tiene en mente, y los economistas procuran investigar modelos de una economía en

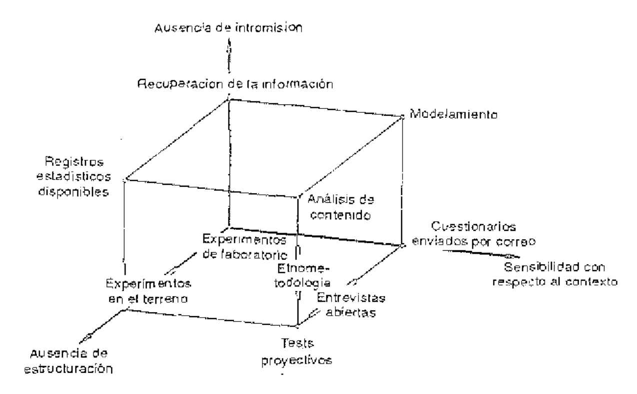


Figura 2. Ambitos empíricos del análisis de contenido y contraste entre las técnicas.

lugar del sistema real que ésta constituye. Además, *el investigador puede querer ocultar su interés por los datos para evitar que la fuente haga de esto un uso instrumental*, ya que las aseveraciones de carácter instrumental son difíciles de analizar. Si Goebbels hubiera sabido de qué manera, mediante qué métodos y con qué fines iban a analizarse sus emisiones radiofónicas durante la segunda guerra mundial, sin duda habría encontrado el modo de engañar al analista.

2. *El análisis de contenido acepta material no estructurado.* Desde luego, la estructuración de una situación de modo que los datos lleguen al investigador en una forma analizable, tiene mucho valor. Los cuestionarios enviados por correo y los programas de entrevistas ofrecen a los sujetos opciones preestablecidas cuyos significados resulta fácil codificar. En los experimentos, virtualmente se enseña a los sujetos un lenguaje de datos que no les es propio, como por ejemplo apretar determinados botones, indicar mediante números los puntos de una escala, identificar formas o figuras, o administrar *shocks* eléctricos a otros sujetos. También la modelización por ordenador y la recuperación de información de un banco de datos resultan preestructuradas, a raíz de las exigencias formales de su manipulación y almacenamiento. Aunque la investigación mediante encuestas, los experimentos psicológicos y las aplicaciones del ordenador deben su éxito a la facilidad para analizar datos preestructurados, puede que el investigador no disponga de esa opción. *Tal vez le interese el material después de que éste haya sido generado por una fuente* que emplea un lenguaje, una lógica y unas categorías que no son necesarias para el análisis ni compatibles con sus exigencias. Los historiadores, los analistas del contenido de los medios de comunicación de masas y los psicólogos que emplean datos biográficos se ven imposibilitados para obtener la información en la forma en que lo desearían, y deben tomarla tal como se les ofrece por razones que nada tienen que ver con el análisis. Por añadidura, *el investigador quizá no pueda anticipar todas las categorías del análisis y las formas de expresión antes de haber obtenido y examinado el material.* Por ejemplo, las comunicaciones interpersonales suelen ser tan complejas y, *a priori*, tan inciertas en su contenido, que al grabarlas mediante videocassettes tal vez sea preciso examinarlas luego repetidamente, a veces imagen por imagen, si el analista quiere comprender lo que realmente sucedió. Por último, *el analista puede considerar que*

unas limitaciones excesivas impuestas a la forma de las respuestas del sujeto pueden obstaculizar la validez de sus datos. Este es particularmente el caso de los análisis del contenido de conceptualizaciones y expresiones espontáneas.

3. *El análisis de contenido es sensible al contexto y, por lo tanto, es capaz de procesar formas simbólicas.* Al analizar las respuestas obtenidas en diversas condiciones experimentales, los datos se disocian de los significados simbólicos que dichas respuestas pueden haber supuesto para los sujetos en cuestión, y se analizan como un conjunto de datos con independencia de sus características simbólicas. Así, aunque los experimentos no impiden el uso del lenguaje, las respuestas verbales no suelen tratarse como fenómenos simbólicos, salvo en lo referente a los significados teóricos que el investigador les atribuye. Esta limitación influye asimismo en los procedimientos de recuperación de información que recurren a las equiparaciones sintácticas y en los criterios de búsqueda que dejan para el usuario del sistema la interpretación simbólica de los datos.

Aunque evidentemente es más fácil reducir las cualidades simbólicas de los datos a etiquetas de categorías, variables y datos pun uales que no tienen consecuencia alguna para el proceso de análisis, y sólo resultan significativos para el investigador que interpreta sus hallazgos en esos términos, hay dos motivos por los cuales un investigador puede preferir hacer exactamente lo contrario. *Un analista puede querer analizar los datos verbales como fenómenos simbólicos*, y en el proceso de transformación conservar la referencia a lo que representan, o causan, o controlan, o constituyen, o reproducen, o a aquello con lo cual están asociados en el interior del contexto original. Así, el analista político que mediante un estudio de caso examina un tratado internacional o un discurso presidencial no puede simplemente pasar por alto las características semánticas y situacionales, ni tampoco las consecuencias políticas de esos sucesos. La dirección de la interpretación del analista es isomorfa con respecto al proceso simbólico en la realidad. O tal vez *desee analizar los datos en relación con un contexto que no comparten los comunicadores o sujetos individuales en cuestión, o que sencillamente no conceptualizan de forma similar*. Las eorías válidas, las construcciones analíticas o las experiencias ocurridas en el pasado con un contexto determinado pueden hacer que los datos se conviertan en "entidades simbólicas no convencionales", llevando al

analista a inferencias de las que los sujetos no son conscientes, o que no pueden aceptar o validar. Esta situación se ve bien reflejada en el uso de los tests proyectivos: en ellos los sujetos responden a estímulos ambiguos sugestivos, y así no saben de qué manera le están informando al analista sobre su personalidad, su psicopatología o sus aptitudes. Sólo el investigador puede reconocer las cualidades simbólicas de estos datos.

4. *El análisis de contenido puede abordar un gran volumen de información.* Gran parte de los estudios etnometodológicos, así como los estudios de caso en psiquiatría, historia y ciencias políticas, se centran en fragmentos pequeños y singulares de textos, tratando de inferir con alguna profundidad los correlatos contextuales. Sin duda, este enfoque puede ser eficaz y no se excluye su uso en el análisis de contenido, especialmente en su etapa preliminar, pero tiende a eludir el uso de controles de calidad para garantizar el significado estadístico y la fiabilidad de los hallazgos. Los datos generados por los análisis de contenido pueden muy pronto exceder la capacidad de trabajo de un solo analista, como pone en evidencia esta serie de unidades de análisis:

481 conversaciones personales (LANDIS y BURTT, 1924).
427 textos escolares (PIERCE, 1930).
4022 lemas publicitarios (SHUMAN, 1937, citado por BERELSON, 1952).
8039 editoriales de periódicos (FOSTER, 1938).
800 noticias difundidas por programas de radio en lenguas extranjeras (ARNHEIM y BAYNE, 1941).
19.553 artículos editoriales (POOL, 1952).
15.000 personajes en mil horas de programas de ficción televisivos (GERBNER y otros, 1979).

En consecuencia, *el especialista en análisis de contenido puede encontrarse ante un gran volumen de datos lingüísticos que ya no podrá analizar una sola persona.* Tal vez deba contar con muchos colaboradores con el auxilio de máquinas; el procesamiento abarcará un período prolongado y exigirá una organización especial de la investigación, controles de calidad y una metodología bien definida.

3. El uso de la inferencia y sus clases

La diferenciación de las diversas formas de inferencias que los analistas pueden utilizar permite distinguir las tareas indispensables en el análisis de contenido. En este capítulo, los mecanismos lógicos empleados para relacionar los datos con su contexto se examinan en función de los sistemas, normas, índices, representaciones lingüísticas, comunicaciones y procesos institucionales.

El análisis de contenido ocupa un lugar importante dentro de la metodología de los instrumentos de investigación. Ante todo, permite aceptar como datos comunicaciones simbólicas comparativamente no estructuradas y, en segundo lugar, permite analizar fenómenos no observados directamente a través de los datos relacionados con ellos, independientemente de que intervenga o no un lenguaje. Dado que la mayoría de los procesos sociales se llevan a cabo a través de símbolos, en las ciencias sociales y en las humanidades es donde se encuentra más difundido el uso del análisis de contenido. Varios autores han identificado y clasificado sus tipos y aplicaciones. JANIS (1965) ofrece la clasificación siguiente:

1. *Análisis de contenido pragmático*: procedimientos que clasifican los signos según su causa o efecto probable (por ejemplo, cómputo de la cantidad de veces que se dice algo que puede producir

como efecto una actitud favorable hacia Alemania en un público determinado).

2. *Análisis de contenido semántico*: procedimientos que clasifican los signos de acuerdo con sus significados (por ejemplo, cómputo de la cantidad de veces que se hace referencia a Alemania, sin importar las palabras específicas que se utilicen en esa referencia).
   a) *Análisis de designaciones*: proporciona la frecuencia con que se hace referencia a determinados objetos (personas, cosas, grupos o conceptos), es decir que es aproximadamente equivalente a un análisis temático (por ejemplo, referencias a la política exterior alemana).
   b) *Análisis de atribuciones*: proporciona la frecuencia con que se remite a ciertas caracterizaciones de un objeto (por ejemplo, referencias a la deshonestidad).
   c) *Análisis de aseveraciones*: proporciona la frecuencia con que ciertos objetos son caracterizados de un modo particular, es decir, equivale aproximadamente a un análisis temático (por ejemplo, referencias a la política exterior alemana como deshonesta).
3. *Análisis de vehículos-signos*: procedimientos que clasifican el contenido de acuerdo con las propiedades psicofísicas de los signos (por ejemplo, cómputo de la cantidad de veces que aparece la palabra "Alemania").

Por su parte, BERELSON (1952) enumera diecisiete aplicaciones del análisis de contenido:

- Para describir tendencias en el contenido de las comunicaciones.
- Para seguir el curso del desarrollo de estudios académicos.
- Para establecer las diferencias internacionales en materia del contenido de las comunicaciones.
- Para comparar los medios o "niveles" de comunicación.
- Para verificar en qué medida el contenido de la comunicación cumple los objetivos.
- Para construir y aplicar normas relativas a las comunicaciones.
- Para colaborar en operaciones técnicas de una investigación (por ejemplo, codificar preguntas abiertas en las entrevistas de encuestas).
- Para exponer las técnicas de la propaganda.
- Para medir la "legibilidad" de los materiales de una comunicación.
- Para poner de relieve rasgos estilísticos.
- Para identificar los propósitos y otras características de los comunicadores.

- Para determinar el estado psicológico de personas o de grupos.
- Para detectar la existencia de propaganda (fundamentalmente con fines legales).
- Para obtener información política y militar.
- Para reflejar actitudes, intereses y valores ("pautas culturales") de ciertos grupos de la población.
- Para revelar el foco de la atención.
- Para describir las respuestas actitudinales y conductuales frente a las comunicaciones.

Stone y Dunphy (STONE y otros, 1966) describen aplicaciones del análisis de contenido en los siguientes campos empíricos:

- Psiquiatría
- Psicología
- Historia
- Antropología
- Educación
- Filología y análisis literario
- Lingüística

Consideran asimismo estos autores que el periodismo y la comunicación de masas constituyeron el origen histórico de esta técnica.

HOLSTI (1969), al igual que Janis, sitúa los datos en el interior del contexto de la comunicación entre un emisor y un receptor, y pasa revista a los análisis del contenido en función de tres finalidades principales:

- Describir las características de la comunicación, averiguando *qué* se dice, *cómo* se dice y a *quién* se dice.
- Formular inferencias en cuanto a los antecedentes de la comunicación, averiguando *por qué* se dice algo.
- Formular inferencias en cuanto a los efectos de la comunicación, averiguando *con qué efecto* se dice algo.

En lo que sigue nos apartaremos de lo anterior, y diferenciaremos los usos de las técnicas en función de las formas de inferencias que pueden realizarse en los análisis de contenido. Estas formas son:

- *Sistemas*
- *Normas*

- *Indices y síntomas*
- *Representaciones lingüísticas*
- *Comunicaciones*
- *Procesos institucionales*

Nuestras distinciones están motivadas por las diferencias entre los mecanismos que aplican los analistas del contenido para relacionar los datos con su contexto, que a su vez implican diferencias en el tipo de conocimiento contextual que intentan obtener, las construcciones analíticas que necesitan y la manera en que pueden validarse los hallazgos. Por supuesto, la eficacia de cualquiera de estas formas depende de la situación empírica.

Nuestro propósito en este apartado es definir estas formas de inferencias e ilustrar cómo han sido utilizadas, sin pretender exponer un panorama completo de este campo.

SISTEMAS

Un sistema es un artificio conceptual que describe una porción de la realidad. Como mínimo, comprende:

- *Componentes* cuyos estados son variables.
- *Relaciones* que se manifiestan en las limitaciones de la co-ocurrencia de estados de los componentes.
- *Transformaciones* de acuerdo con las cuales ciertas relaciones implican a otras en el tiempo o en el espacio.

Los sistemas permiten extrapolar los datos existentes a otros estados de cosas aún desconocidos, y en este sentido ofrecen explicaciones autónomas. El sistema solar puede describirse en estos términos: la configuración de los planetas sigue una secuencia temporal definida; para alguien que conozca el sistema, los datos de una configuración cualquiera contienen implícitamente los de todas las configuraciones siguientes. También la terminología del parentesco constituye un sistema en este sentido. Las relaciones entre individuos están definidas en su interior por las semejanzas de descendencia, matrimonio, adopción y diferencia entre los sexos, y un código prescribe de qué manera se han de designar unos a otros los individuos emparentados. Un sistema permite formular extrapolaciones referidas a cualquier individuo

nuevo que ingrese en una relación. Los sistemas de entidades biológicas, simbólicas o sociales pueden llegar a ser muy complicados, pero todos ellos retienen la idea básica de un eslabonamiento temporal o espacial de relaciones entre sus numerosos componentes. Las inferencias que revisten interés para el análisis de contenido proceden de transformaciones invariantes dentro de un sistema de símbolos, y pueden extenderse más allá del tiempo y el espacio de los datos disponibles.

La idea de un enfoque sistémico se remonta a TENNEY (1912), quien se preguntó: "¿Por qué la sociedad no puede estudiar sus propios métodos de producción de las diversas variedades de pensamiento, estableciendo .. un cuidadoso sistema de registro?... Lo que se necesita... es el análisis permanente de un gran número de periódicos... los registros en sí mismos constituirían una serie de observaciones del 'clima social', comparable por su exactitud con las estadísticas del Servicio Meteorológico de los Estados Unidos". Tenney halló ciertas interacciones sistemáticas entre los diversos temas de que se ocupaban los periódicos, sospechó determinados cambios en dichas relaciones a lo largo del tiempo, y exploró en especial las características étnicas de esas publicaciones. Equiparó esta dinámica de la cobertura de prensa en todo un país con el proceso de pensamiento de ese país, pero le faltaron datos y, quizá también, buenos métodos para describirlo en forma sistemática.

En época más reciente, RAPOPORT (1969) sentó las bases de una teoría sistémica de los "corpus textuales", planteando, entre otros interrogantes, qué significa describir un gran "corpus" de datos simbólicos diciendo que tiene un determinado comportamiento, que cambia y evoluciona, así como el carácter de los componentes, relaciones y leyes de interacción que resultarían adecuados en los corpus de esa índole. Aunque Rapoport tiene conciencia de que el medio semántico en que se desenvuelve el hombre refleja y a la vez constituye gran parte de su existencia, de que se ve continuamente enriquecido y quizá también contaminado por actos individuales y políticas institucionales, sugiere que el estudio a gran escala de los productos textuales de esta índole puede ser más fructífero si se efectúa sin referencia al usuario del símbolo, es decir (al menos por algún tiempo), como sistemas autónomos.

Tendencias

El prototipo de un enfoque sistémico en el análisis de contenido es la extrapolación de tendencias. SPEED (1893) comparó varios periódicos neoyorquinos publicados en 1881 con los que se publicaban doce años después, observando variaciones en las categorías temáticas. Aunque observaciones efectuadas en sólo dos momentos difícilmente puedan prestarse a la realización de predicciones sólidas, el análisis de Speed demostró con toda claridad las intenciones dominantes, lamentando la disminución de la crítica literaria y el aumento de la "chismografía", de los artículos deportivos y la ficción. LASSWELL (1941) propuso estudiar las tendencias establecidas en la frecuencia con que se mencionaban diversos países en la prensa nacional de otras tantas naciones, presentando hallazgos preliminares. LOEVENTHAL (1944) examinó la variable definición de los "héroes populares" de las revistas, demostrando que se tendía a escoger cada vez menos a los héroes entre los profesionales y los hombres de negocios, y a encarnarlos cada vez más en personas relacionadas con la industria del espectáculo, tendencia que aún sigue vigente. Otros estudios sobre las tendencias se ocuparon de los valores presentes en la literatura de inspiración religiosa, la publicidad, los *slogans* políticos y la frecuencia de las publicaciones en determinadas disciplinas científicas. Uno de los más amplios análisis de contenido que utilizaron esta perspectiva es el estudio de NAMENWIRTH (1973) sobre los cambios de valores manifiestos en las plataformas de los partidos políticos norteamericanos a lo largo de un período de 120 años. Descubrió que en estos datos existían dos ciclos independientes, capaces de explicar los cambios globales en los valores como un sistema autónomo.

Pautas

Otra noción sistémica presente en el análisis de contenido es el uso predictivo de las pautas. En el folclore, el análisis estructural de los acertijos, proverbios, leyendas y relatos aborda específicamente la identificación de pautas que tengan un alto grado de predecibilidad dentro de un género, con independencia de los contenidos particulares (ARMSTRONG, 1959). Dicho análisis parte de discernir los principales elementos constitutivos del género, y procura dilucidar la lógica que relaciona entre sí esos elementos.

Así, SEBEOK y ORZACK (1953), en su análisis de los encantamientos de cheremis, comprobaron que a "una declaración puramente fáctica sobre el mundo" le seguía "un motivo cuya ocurrencia era extremadamente improbable". El análisis de relatos de LABOV (1972) se inserta dentro de una tradición similar. Otro ejemplo es el análisis de la genealogía dentro de un corpus literario, a través de las pautas de citas. Las publicaciones científicas tienden a citar otras publicaciones, entrelazándose así mutuamente en una red bibliográfica. GARFIELD (1979) recurre a este hecho para elaborar su índice de citas. En la investigación de las comunicaciones, el establecimiento de "quién dice qué cosa a quién" permite trazar diferentes pautas, redes de comunicación que se asemejan a los sociogramas; asimismo, el estudio de las co-ocurrencias dentro de una oración, un párrafo o todo un documento conduce a pautas de co-ocurrencias a partir de las cuales pueden extraerse inferencias sobre un corpus literario.

La combinación de las *tendencias* con las *pautas* ha dado origen a numerosos e interesantes análisis de contenido. El *análisis del proceso de interacción*, de BALES (1950), estableció pautas de comunicación, evaluación, control, toma de decisiones, reducción de la atención y reintegración, todo lo cual se definió a partir de doce categorías básicas de intercambios verbales en pequeños grupos. HOLSTI y otros (1965) analizaron la sucesión de declaraciones públicas efectuadas durante la crisis de los misiles cubanos en 1962 por importantes funcionarios del gobierno norteamericano y de la Unión Soviética, clasificándolas en percepciones y expresiones, y describiendo luego cada una de éstas en función de los valores de la dimensión evaluativa, de intensidad y de potencia, dimensiones tomadas de la teoría de las actitudes de Osgood. Estos autores obtuvieron de ese modo datos para un modelo de interdependencia dinámica que demostró ser bastante predictivo con respecto a las respuestas emocionales que cada uno de estos grupos de funcionarios daba al otro.

Diferencias

Todos los enfoques sistémicos se ocupan de extrapolar diferencias a nuevas situaciones. De hecho, gran cantidad de obras sobre el análisis de contenido evalúan las diferencias existentes en los mensajes generados por dos comunicadores o por una misma fuente en dos situaciones distintas, o las que surgen de acuer-

do con el público al que están dirigidos los mensajes, o entre los datos de entrada [*input*] y los datos de salida [*output*]. Las diferencias en la cobertura periodística de las campañas políticas se han relacionado con el apoyo otorgado a esas campañas en artículos editoriales (KLEIN y MACCOBY, 1954). Las diferencias en la cobertura periodística de las manifestaciones sobre derechos civiles han sido explicadas en función de diversas características de los periódicos, incluido el lugar geográfico en que se editan, quiénes son sus dueños y cuál es su orientación política (BROOM y REECE, 1955).

También se estudiaron las diferencias entre periódicos rivales o con un monopolio de la información (NIXON y JONES, 1956). En la crónica de un crimen apolítico por parte de la prensa francesa (GERBNER, 1964) se puso de relieve hasta qué punto difieren las orientaciones ideológicas y de clase en el flujo de los mensajes. Los mensajes procedentes de una fuente pueden variar, asimismo, según el público al que estén dirigidos, como se demostró comparando los discursos políticos de John Foster Dulles ante grupos de diferente composición (COHEN, 1957; HOLSTI, 1962). También las obras de ficción escritas difieren según estén destinadas a la clase alta, media o baja (ALBRECHT, 1956), y lo mismo sucede con la publicidad en las revistas según sus lectores sean blancos o negros (BERKMAN, 1963). Como ejemplo de las diferencias entre los datos de entrada y de salida se pueden citar los estudios que relacionan las fuentes de información de un periódico con lo que finalmente aparece impreso (ALLPORT, 1940), los análisis de lo que pasa con un libro cuando se adapta para el cine (ASHEIM, 1950) o la comparación de los hallazgos científicos con sus versiones de divulgación popular.

Un buen ejemplo de combinación de las *diferencias entre los medios* y las *tendencias* es el estudio llevado a cabo por la Hoover Institution sobre la Revolución y el Desarrollo de las Relaciones Internacionales (cuya sigla inglesa es RADIR*). En él se identificaron símbolos clave de la "democracia", la "igualdad", los "derechos" y la "libertad" en 19.553 artículos editoriales aparecidos en prestigiosos periódicos norteamericanos, británicos, franceses, alemanes y rusos en el período 1890-1949. El análisis de estos datos llevó a POOL (1951) a observar-predecir que las doctrinas proletarias fueron reemplazando a las tradiciones libe-

* Revolution and the Development of International Relations. [E.]

rales, que una creciente amenaza bélica estaba relacionada con el aumento del militarismo y el nacionalismo, y que la agresividad hacia otras naciones se vinculaba con la inseguridad. Aunque todos estos símbolos aluden a diversos aspectos de la realidad política, fueron categorizados, computados y analizados por un especialista sin tener en cuenta lo que representaban. La observación de Pool acerca de que los símbolos de la democracia aparecen con menos frecuencia cuando una forma representativa de gobierno goza de aceptación que cuando está en tela de juicio, sugeriría, empero, que este sistema simbólico no es del todo autónomo. El estudio de GERBNER y otros (1979) sobre la violencia televisiva en los programas de ficción ha acumulado ya suficientes datos como para formular extrapolaciones de interés para las autoridades encargadas de establecer las políticas en esta materia. Su *análisis sistémico del mensaje* (GERBNER, 1969) procura hacer esto en lo referente a cuatro áreas generales de interés:

1. La frecuencia con que aparecen los componentes de un sistema, o sea, "lo que es".
2. El orden de prioridades asignado a esos componentes, o sea, "lo que es importante".
3. Las cualidades afectivas asociadas con los componentes, o sea, "lo que es correcto".
4. Las asociaciones próximas o lógicas entre los componentes, o sea, "qué se relaciona con qué".

Lamentablemente, la mayoría de las aplicaciones prácticas de las nociones sistémicas al análisis de contenido se han visto obstaculizadas por sus formulaciones todavía arcaicas. A menudo, el estudio de las tendencias sólo se ha ocupado de una variable o ha tomado una sola cada vez (como las frecuencias, el volumen, el crecimiento, la definición o los valores). Por su parte, las pautas se han conceptualizado a menudo en relaciones binarias, como las de sucesión, contraste, proximidad, causa o comunicación. Y rara vez se han explicado las diferencias en función de las interacciones que las realzan o las atenúan con el transcurso del tiempo. Uno de los problemas que se plantean es el volumen de datos necesarios para descubrir características suficientemente invariantes que permitan extraer inferencias sistémicas. A medida que este problema comenzó a tratarse a través de los medios electrónicos de procesamiento de información, pasó a primer plano otro: ¿cómo nos es posible desarrollar concepciones sistémicas ade-

cuadas, que puedan prestarse al cálculo de los datos simbólicos como sistemas por derecho propio?

NORMAS

Los procesos de identificación, evaluación y verificación tienen en común la existencia de una norma o patrón con la que se compara un objeto para establecer de qué clase es o en qué medida es bueno. Aunque a primera vista parezca que estos procesos no suponen inferencias, lo cierto es que las identidades no se revelan en el vacío. Toda evaluación carece de significado si no tiene un propósito, y las verificaciones son intrascendentes si los criterios aplicados no gozan de apoyo institucional. El resultado de estos procesos depende de la interacción entre las características de un objeto y la norma, varía con diferentes normas y, por lo tanto, implica inferencias, aunque sean de una especie simple.

Evaluaciones

La evaluación de la actuación de los medios de prensa ha sido una preocupación fundamental desde que apareció el análisis cuantitativo de periódicos. El interés por el cambio de la calidad a la cantidad en la información (SPEED, 1893) o por el aumento de los temas triviales, desmoralizadores o malsanos a expensas de los dignos (MATTHEWS, 1910) da por sentadas ciertas normas de evaluación, aunque sea de forma implícita. Aunque es posible compartir la inquietud que condujo a la realización de estos estudios, es difícil ponerse de acuerdo en una escala aceptable que sitúe a los periódicos en una gama que vaya de lo bueno a lo malo. Muchos de los estudios evaluativos se han limitado a medir la orientación hacia uno u otro bando en una controversia; por ejemplo, ASH (1948) intentó determinar si al público se le habían dado buenas oportunidades de informarse acerca de lo que pensaban ambos bandos en la controversia sobre la ley Taft-Hartley de reglamentación del derecho de huelga.

La mayoría de los estudios evaluativos carecen de criterios defendibles. JANIS y FADNER (1965) publicaron un "coeficiente de desequilibrio" que mide el grado en que las declaraciones favorables superan en número a las desfavorables. Formalmente, el coeficiente es impecable; pero sigue en pie la cuestión de si los

periodistas deben ser imparciales o en qué circunstancias deben serlo. Por ejemplo, en los últimos días de la presidencia de Nixon era difícil no tomar partido. MERRILL (1962) utilizó una batería de criterios de evaluación de presentaciones periodísticas (orientación de atribución, orientación adjetiva, orientación adverbial, orientación contextual, orientación fotográfica y opinión franca), pero su catálogo dista mucho de ser completo. BERELSON y SALTER (1946) emplearon la composición racial de la población norteamericana como norma para evaluar la población de personajes de ficción en las revistas populares, idea que dio origen a muchos otros estudios evaluativos (BERKMAN, 1963). Pero tampoco resulta claro si la población de los personajes de ficción en las revistas, el teatro o la televisión puede ser estadísticamente representativa de ciertas características del público, como la etnicidad, el sexo, la edad y, en especial, las actividades a las que se dedica. El Council on Inter-racial Books for Children (1977) propuso y expuso un método para evaluar los textos de historia de los Estados Unidos comparándolos con los "datos históricos". No obstante, probablemente sea inevitable cierto grado de selectividad en la presentación de la historia. Lo que provoca inquietud, y debería ser el eje de las decisiones en todo trabajo evaluativo, son los excesos sistemáticos, no la adhesión a una norma fija.

Identificaciones

Si las evaluaciones establecen en qué grado algo se adecua o se aparta de una norma, las identificaciones tienen un carácter más disyuntivo, del tipo "o bien... o bien...". DIBBLE (1963) consideró la clase de inferencias que interesan a los historiadores, ejemplificando de la siguiente manera su categoría de "documentos como indicadores directos": supóngase que un historiador desea saber si el embajador británico en Berlín se comunicó con el ministro de relaciones exteriores de su país el día anterior al estallido de la primera guerra mundial. La carta que figura en los archivos del ministro ofrece una prueba directa de que ha sido enviada. La inferencia del investigador le exigirá identificar dicha carta mediante el cumplimiento de ciertos requisitos. Las identificaciones de esta índole incluyen la autentificación de obras de arte, la verificación de un testamento o la certificación de una firma. Incluso en un nivel "micro", las identificaciones son elementos importantes en el análisis de contenido mediante

ordenador. Por ejemplo, el programa llamado General Inquirer (STONE y otros, 1966) clasifica las palabras de un texto en clases de palabras semánticamente semejantes. Pero como el programa no reconoce la similitud de significado, esta idea debe incorporarse a un "diccionario" que reduce la tarea de clasificación a la de "identificación de características especificadas". Gran parte del análisis de contenido supone la categorización de vehículos-signos mediante el discernimiento de los significados que transmiten a un público o a un receptor determinados (JANIS, 1965).

Verificaciones

También las verificaciones suponen emitir juicios sobre los datos con respecto a una norma, con el añadido de que esta norma es *prescrita* o legitimada por una institución. Por ejemplo, cuando la FCC exige que todos los canales de televisión emitan una cierta proporción de programas con y sin patrocinio comercial, así como programas de actualidad local y de interés público, ello implica una forma de medición y la aplicación de determinados criterios. Las decisiones sobre el plagio, la obscenidad, las prácticas discriminatorias, e incluso sobre el hecho de si un programa de ficción puede estimular el delito, podrían basarse en pruebas extraídas a partir del análisis de contenido, siempre y cuando la ley fuera lo suficientemente concreta en cuanto a la naturaleza de las pruebas admisibles. Lamentablemente, en su mayoría, las cuestiones legales correspondientes al contenido de las comunicaciones no se formulan en un lenguaje fácilmente operativo. Con todo, los resultados de los análisis de contenido han sido aceptados en los tribunales, al menos desde que Lasswell (1949b) emprendiera su análisis comparativo de las publicaciones de ciertos individuos sospechosos de ser agentes de países extranjeros (véase su libro *Content Analysis*, 1948).

INDICES Y SINTOMAS

Un *índice* es una variable cuya importancia en una investigación depende del grado en que pueda considerarse *correlato de otros fenómenos*. Pierce estableció el requisito de que todo índice debe estar causalmente relacionado con el suceso que significa (del mismo modo que el humo es índice del fuego) o basado en

una necesidad física o material, y no en una convención arbitraria (símbolo) ni en la similitud (icono). En medicina, a los índices se los denomina *síntomas*. Muchas afecciones y trastornos biológicos demuestran traer consigo efectos colaterales visibles, que el paciente no puede controlar a voluntad. Una parte significativa de los diagnósticos médicos incluye la interpretación de los cambios visibles en el cuerpo humano como símbolos de una enfermedad. "Un índice... no depende de las entidades o sucesos físicos de los que procede" (RAPOPORT, 1969).

Análogamente, en muchos análisis de contenido se utilizan entidades medibles como índices de otros fenómenos que no pueden medirse de forma tan directa. Típicamente, los datos lingüísticos (o paralingüísticos) se miden de modo tal que adquieren el carácter de índices de fenómenos no lingüísticos: el cociente de perturbación en el habla mide la angustia del paciente durante una entrevista psiquiátrica (MAHL, 1959); la frecuencia del empleo de cierta clase de palabras es un índice de la motivación de rendimiento (GERBNER y otros, 1979); se han creado índices de la insatisfacción de los ciudadanos medida a través de las cartas en que manifestaron sus quejas (KRENDEL, 1970), así como el criterio de legibilidad de FLESH (1948-1951); la relación adjetivos/verbos fue utilizada por BRODER (1940) como índice de la esquizofrenia (MANN, 1944); y las co-ocurrencias improbables de nombres, como indicadores de asociaciones similares en el hablante y el receptor (OSGOOD, 1959).

En las investigaciones sobre la comunicación de masas, hay tres índices cuyo uso es de origen antiguo:

- La *frecuencia* con que aparece un símbolo, idea o tema en el interior de una corriente de mensajes tiende a interpretarse como medida de *importancia, atención* o *énfasis*.
- El equilibrio en la cantidad de atributos *favorables y desfavorables* de un símbolo, idea o tema tiende a interponerse como medida de la *orientación o tendencia*.
- La cantidad de asociaciones y de calificaciones manifestadas respecto de un símbolo, idea o tema suele interpretarse como una medida de la *intensidad o fuerza* de una creencia, convicción o motivación.

El empleo de cantidades fácilmente computables en calidad de índices no deja de presentar dificultades. CHOMSKY (1959), opo-

niéndose a la sugerencia de Skinner acerca de que la rapidez de respuesta, la cantidad de repeticiones, el volumen de la voz, etc., son índices naturales de la intensidad de la motivación de un sujeto, compara a dos mujeres hipotéticas, cada una de las cuales recibe un enorme ramo de flores. La primera, en cuanto ve las flores, comienza a gritar con toda la fuerza de sus pulmones: "¡Hermosas! ¡Hermosas! ¡Hermosas! ¡Hermosas!", poniendo en evidencia así, de acuerdo con el criterio de Skinner, que posee una fuerte motivación para producir esa respuesta. La segunda abre el paquete y, una vez vistas las flores, no dice nada durante diez segundos, para luego sólo susurrar, con voz apenas audible: "Hermosas". Presumiblemente, esta última exhibiría una "débil motivación para la respuesta" (RAPOPORT, 1969).

También resultan problemáticos los índices de atención que se basan en la frecuencia. Una cosa es utilizar las frecuencias o repeticiones para otorgar certidumbre a una hipótesis, y otra muy distinta utilizarlas como indicadores de un fenómeno que debe correlacionarse con ellas. Lo primero es un procedimiento científico, lo segundo una propiedad empírica. LASSWELL (1965a), en el famoso artículo en que se pregunta "¿Por qué ser cuantitativo?" en el análisis de contenido, confunde esta distinción. Las medidas de frecuencia pueden constituir indicadores eficaces si el fenómeno subyacente está asimismo relacionado con una frecuencia. Por ejemplo, es probable que la cantidad y calidad de las cartas en que se expresan quejas ante el ayuntamiento (KRENDEL, 1970) indiquen la misma insatisfacción que se manifestaría en una elección de alcalde. En un momento crítico, BERELSON (1952) se preguntaba qué podría inferir un marciano al comprobar, en los medios de comunicación de masas que sobrevivieran a nuestra época, la alta frecuencia de los temas de amor y sexo: ¿que vivíamos en una sociedad promiscua o en una sociedad represiva? POOL (1952b) observó que los símbolos de la democracia aparecen con menor frecuencia allí donde se aceptan los procedimientos democráticos que allí donde son objeto de discusión.

Las dificultades que entraña la creación de un índice se ponen de manifiesto en el uso del *cociente de alivio de la incomodidad*. Basándose en consideraciones extraídas de la teoría del aprendizaje, DOLLARD y MOWRER (1947) interpretaron dicho cociente como índice de medición de la angustia de un hablante. Se calculaba como una proporción entre la cantidad de "palabras provo-

cadoras de incomodidad" o "palabras-impulso" [*drive-words*] y la suma de la cantidad de palabras de incomodidad o palabras-impulso más la cantidad de "palabras que producen comodidad o alivio". A pesar de los argumentos teóricos en su favor, las pruebas sobre la eficacia de este cociente como índice han dado origen a resultados heterogéneos. Se ha informado de correlaciones significativas con el sudor de la palma de la mano, pero en cambio las correlaciones con otras mediciones de la angustia sólo parecen demostrables en circunstancias muy limitadas. MURRAY y otros (1954) compararon el *cociente de alivio de la incomodidad* con otras diversas medidas de la motivación y el conflicto utilizadas en el curso de la terapia, y comprobaron que aquél no era sensible a las variaciones en el avance terapéutico. Así pues, no resulta en absoluto claro qué puede indicar dicho cociente.

Algo mayor fue el éxito alcanzado en el uso de los índices de legibilidad. A todas luces, el empleo de expresiones extranjeras, de palabras largas y compuestas, y de una gramática y una puntuación complicadas convierten la lectura en algo difícil. FLESH (1948) reunió en un coeficiente único (su "criterio de legibilidad") las proporciones de todos los pares (por ejemplo, la de palabras extranjeras con respecto a las nativas, y de las palabras largas con respecto a las breves), criterio relacionado con la facilidad de lectura, tal como lo exponían los propios lectores. Estudios subsiguientes consideraron una cantidad mucho mayor de mediciones y diferenciaron factores como el interés humano, la comprensión y la velocidad de lectura, obteniendo así medidas que están mejor relacionadas con el criterio "original". Podría sospecharse que la razón de este éxito, por moderado que sea, radica en una idea añadida acerca de la facilidad de lectura: presumiblemente, cada dificultad con que se topa un lector se añade a su juicio global de la dificultad del texto, más o menos como ocurre cuando se computan palabras difíciles. Esto puede que no sea válido para otras magnitudes.

Los índices pueden ser eficaces, asimismo, cuando se trata de zanjar disputas acerca de la autoría de un texto. YULE (1944), un especialista en estadísticas de seguros, se replanteó si el autor de la *Imitación de Cristo* era Tomás de Kempis, Jean Charlier de Gerson, o algún otro de los varios autores hipotéticos. Correlacionó las frecuencias de los sustantivos que aparecían en las obras que sin duda escribieron algunos de estos autores potenciales, con lo cual pudo establecer índices discriminativos de su

respectiva identidad, para aplicarlos luego a la obra en cuestión. La diferencia fue favorable a Tomás de Kempis. MOSTELLER y WALLACE (1964) comprobaron que las palabras funcionales eran más discriminadoras que los sustantivos para decidir la cuestionada autoría de los doce *Federalist Papers*; la evidencia favoreció a Madison.

Pero el establecimiento de índices no es un asunto que dependa de una definición verbal. El hecho de computar frecuencias y de considerarlas una medida de la atención no supone la creación de un índice de la atención. Aun cuando se empleen datos empíricos en la creación del índice, queda en pie el problema de la generalización. Por ejemplo, MORTON y LEVINSON (1966) analizaron textos griegos de autores conocidos y extrajeron siete marcas de estilo que, según ellos, singularizaban las características de la escritura de cualquier individuo: longitud de la oración, frecuencia de uso del artículo definido y de las palabras "y", "pero" y "en", empleo de los pronombres de tercera persona y de todas las formas del verbo "ser". Tras analizar las catorce epístolas atribuidas a Pablo, MORTON (1963) llegó a la conclusión de que habían sido escritas por seis autores distintos, y que en realidad Pablo sólo era el autor de cuatro. ELLISON (1965) aplicó estas construcciones a textos de autores conocidos, llegando a la conclusión de que el *Ulises*, de James Joyce, había sido escrito por cinco autores distintos, y que ninguno de ellos era el autor de *Retrato del artista adolescente*. Y demostró que incluso el propio artículo de Morton estaba escrito en varios "estilos" distintos, lo cual arrojó serias dudas sobre la utilidad de los índices estilísticos de Morton para determinar la identidad de un autor.

REPRESENTACIONES LINGÜISTICAS

En todo discurso interviene el lenguaje mediante la exposición y argumentación sistemáticas, incluido el examen metódico de los hechos y principios en cuestión y las conclusiones alcanzadas. Un discurso se ocupa de una porción limitada de la realidad, o de alguno de sus rasgos experienciales. Puede tener su origen en una sola persona o en un grupo de personas que interactúan; puede definir su propio asunto o tema, permanecer abierto a la introducción de nuevos hechos y aceptar la modificación de los que anteriormente se creían verdaderos (o así los consideraban

otros individuos). Analizar un corpus textual como discurso implica establecer las relaciones entre dos o más oraciones, siempre y cuando estas oraciones estén vinculadas al conocimiento de la realidad que dicho corpus representa.

HAYS (1969) proporciona los siguientes ejemplos típicos de conjuntos de datos lingüísticos que pueden interesar a los científicos sociales:

- *Secuencias de artículos editoriales.* La plantilla de un periódico, que vive en una determinada época histórica, produce una serie de ensayos en los que recapitula algunos de los sucesos del momento, y los sitúa respecto a las tendencias históricas, las teorías y los dogmas. Expresa sus opiniones acerca de la verdadera naturaleza de situaciones que no deben necesariamente comprenderse del todo, así como sus opiniones acerca de las respuestas suscitadas.
- *Intercambios internacionales de carácter oficial.* Esta clase de correspondencia puede compararse con la secuencia de artículos editoriales periodísticos, salvo que en este caso hay dos o más bandos, cada uno de los cuales persigue una política propia.
- *Documentos personales.* Puede tratarse de cartas o diarios íntimos, o de material escrito de alguna otra especie. Excepto por la particularidad de su contenido, este material puede también compararse con los editoriales periodísticos o el intercambio de documentos oficiales.
- *Transcripciones de entrevistas.* Aquí se trata por lo general de dos bandos, uno de los cuales es el "ingenuo" o desconocedor, mientras que el otro es el "experto", conocedor o entendido. La finalidad de la entrevista puede ser; por ejemplo, terapéutica o diagnóstica.
- *Interacción social.* Dos o más personas participan en el análisis de una tarea prefijada o de cualquier tema que estimen conveniente tratar.

Aun cuando se pase por alto el aspecto secuencial de estos ejemplos —relatos periodísticos de un acontecimiento, testimonios de testigos en los tribunales, informes sobre los descubrimientos de una investigación, tratados académicos y documentos políticos—, lo cierto es que todos ellos tienen en común el uso del lenguaje para representar una porción de la realidad, excluyendo idealmente los sentimientos, intereses y perspectivas subjetivas de la fuente. También la noción intuitiva de "contenido" parece apuntar a una representación lingüística de hechos y experiencias. No obstante, sorprendentemente, los especialistas en

análisis de contenido parecen sentirse más cómodos empleando índices que, aunque se calculan a partir de datos lingüísticos, no se basan en el poder del lenguaje, haciendo en realidad caso omiso de éste. Aunque podrían existir buenas razones para que los análisis de contenido procesaran datos por vías distintas de aquellas mediante los cuales fueron generados, un hecho básico de la comunicación humana es que el lenguaje se utiliza para transmitir conocimientos, y para comprender los conocimientos transmitidos por otros. Por consiguiente, una de las tareas del análisis de contenido es alcanzar esta clase de comprensión de los datos lingüísticos y extraer inferencias sobre esa base.

La forma más rudimentaria de comprender lo que transmite el lenguaje exige clasificar las palabras o expresiones lingüísticas por las referencias (denotaciones, connotaciones) que establecen. Considérese, a título de ejemplo, las numerosas maneras en que se puede hacer referencia en un órgano de prensa al presidente de los Estados Unidos: por su título oficial, por su nombre de pila, por el número que le corresponde dentro de la serie de los presidentes norteamericanos, por la cantidad de años que hace que está en el cargo, por su lugar de residencia (la Casa Blanca) o por sus principales realizaciones políticas. Se requiere un considerable conocimiento lingüístico y político para aislar estas referencias de sus respectivos contextos y situarlas dentro de una misma categoría: la del ocupante de ese cargo. Aunque las diferencias entre expresiones sinónimas desde el punto de vista referencial podrían indicar actitudes o relaciones personales (por mencionar sólo dos factores), cuando se enfoca el contenido de un discurso está justificado hacer caso omiso de esas diferencias. Al cuantificar "la cantidad de atención que se presta a la violencia televisiva", deben discernirse dos componentes: en primer lugar, el *índice* cuantitativo de la magnitud de la atención prestada, y en segundo lugar, la distinción cualitativa entre las *referencias* a las descripciones de escenas de violencia y de no violencia. Este último uso es el que nos interesa aquí y no está relacionado con la cuantificación.

Como es obvio, una simple colección de referencias no hace justicia a todo el saber que es capaz de transmitir un lenguaje. Uno de los usos de las representaciones lingüísticas en el análisis de contenido es el desarrollo de *mapas* cognitivos. GERBNER y MARVANJI (1977) trazaron los mapas del mundo tal como lo verían periódicos norteamericanos, de Europa oriental, de Europa

occidental y de la Unión Soviética, así como algunos periódicos del tercer mundo, utilizando como criterio de magnitud la cantidad de noticias extranjeras. LYNCH (1965), aunque proviene de una tradición distinta, también traza el mapa de todas las enunciaciones verbales acerca del desplazamiento interno dentro de una ciudad, dibujando así el mapa de esa ciudad tal como lo ven sus habitantes. ALLPORT (1965), al analizar las *Cartas de Jenny*, de carácter personal, muestra cómo podría ser el mundo del autor de dichas cartas y qué clase de inferencias psicológicas podrían extraerse de ello. Pueden encontrarse otros ejemplos marginales de representaciones discursivas en las simulaciones de los procesos cognitivos (ABELSON, 1968) y en su aplicación a campañas políticas (POOL y otros, 1964). Tal vez porque los analistas humanos son tan eficientes para la interpretación de las aseveraciones lingüísticas en forma referencial, no existen técnicas detalladas para extraer inferencias de representaciones discursivas que estén tan bien formalizadas.

Veamos las clases de inferencias que podrían obtenerse de esas representaciones. ALLEN (1963), al propugnar una cierta especie de análisis del contenido lógico, muestra cuáles son las opciones disponibles para los signatarios de un tratado de limitación de armamentos; y esto lo lleva a inferir la dirección en que los bandos que firman el tratado pueden o podrían tratar de avanzar y qué conflictos posteriores podrían surgir entre ellos. NEWELL y SIMON (1956), poniendo más el énfasis en las limitaciones que en las opciones disponibles, muestran con su Máquina de la Teoría Lógica de qué manera una secuencia de implicaciones lógicas (una prueba) a partir de los datos disponibles (premisas, axiomas) puede servir para tomar decisiones dentro de un área problemática (la validez de un teorema). Naturalmente, los documentos políticos, los relatos periodísticos, los hallazgos de investigación, etc., tienen interés fundamentalmente a raíz de las *implicaciones* que se extraen de ellos en relación con los problemas que interesan al usuario.

El mejor modelo de un análisis de contenido que utiliza representaciones discursivas se encuentra en las *máquinas de preguntas y respuestas*. Estas vienen regidas por programas de computación, aceptan datos en lenguaje natural (con algunas limitaciones) y pueden responder a preguntas específicas cerciorándose de lo que "dicen" implícitamente los datos disponibles acerca de la pregunta. Este tipo de máquinas poco tiene que ver con los meca-

nismos de recuperación de información, que responden a una pregunta recuperando una idea explícitamente almacenada de acuerdo con algún índice. Tampoco debe confundírselas con los mecanismos para la elaboración de resúmenes que extraen las palabras clave o la oración más representativa de un texto determinado. El rasgo principal de un análisis de contenido de este tipo es que se basa en un mapa del territorio del discurso. Incluye el objetivo del analista en lo que se refiere a los datos lingüísticos ingresados y a las implicaciones personales que se pueden examinar.

Teniendo en cuenta la contribución al análisis de contenido de los recientes avances en lingüística computacional, HAYS (1969) propone un modelo de esta índole denominándolo "conversacionalista". Un análisis de contenido de esta clase debería tener en cuenta lo que se presume que saben los participantes en un discurso. Aceptaría un flujo de datos lingüísticos (el discurso, un diálogo, un tratado, etc.) y respondería a las preguntas del analista en lo referente a las implicaciones, intenciones, discrepancias u opciones.

COMUNICACIONES

Una comunicación es todo mensaje intercambiado entre interlocutores. La composición y contenido de estos mensajes procede, hasta cierto punto, de la intención que persigue su creador, y puede tener diversas consecuencias. Lo más importante es que las comunicaciones se intercambian dentro del contexto de las relaciones vigentes entre los comunicadores, modificando en el proceso esas relaciones. En un primer nivel, las comunicaciones pueden contribuir a explicar causas y efectos entre los cuales existe una mediación simbólica. Además, pueden servir para explicar la dinámica del comportamiento, las consecuencias individuales o colectivas del intercambio de información, diversas psicopatologías, la aparición del conflicto y del consenso y la transformación de una cultura material. Pese a que esta concepción de las comunicaciones parece obvia, son bastante infrecuentes, en realidad, los análisis de contenido que la emplean. A todas luces, los intercambios entre seres dotados de propósitos trascienden las simples representaciones lingüísticas. Aunque el afán de estudiar la interacción a través de los mensajes es sin duda anterior a los análisis del contenido, la formulación de concepciones

adecuadas para extraer las correspondientes inferencias es relativamente reciente.

En su intento de inferir el grado de angustia a partir del registro grabado del habla, MAHL (1959) examinó el modo en que las intenciones de los hablantes afectaban su elección de las palabras. Aunque al final resolvió este problema desarrollando índices adecuados, que no son controlables de forma consciente, su examen del uso instrumental del lenguaje puso de relieve una forma de significado que tiene sus raíces en la circularidad esencial de los efectos de la comunicación: toda aseveración es instrumental en la medida en que estimula una realimentación deseada por el comunicador, independientemente de sus contenidos semánticos. La interpretación de datos lingüísticos como comunicaciones permite al analista formular inferencias acerca de una pauta causal circular.

PROCESOS INSTITUCIONALES

Además de ser indicativos, de transmitir contenido o de poseer consecuencias instrumentales o no deseadas, los mensajes pueden desempeñar también funciones dentro de las organizaciones y las instituciones sociales. Se ha equiparado la comunicación con el aglutinante que mantiene unidas las organizaciones sociales. De hecho, la existencia de familias, organismos de gobierno y sociedades es impensable sin formas regulares y normales de procesos de comunicación en curso. Dentro de las organizaciones, los mensajes pueden cumplir muchos propósitos, de los cuales el más obvio es el de suministrarles información exterior a ellas. Los mensajes son la espina dorsal simbólica de cualquier organización viva, y los análisis de contenido pueden tener como finalidad inferir las estructuras y los procesos institucionales a que dan lugar los datos con que se cuenta.

Lasswell (1960), por ejemplo, distingue tres funciones sociales principales de las comunicaciones, partiendo de las cuales un enfoque institucional del análisis de contenido puede tratar de formular inferencias:

1) La vigilancia del medio ambiente.
2) La correlación entre las distintas partes de la sociedad en su reacción frente al medio ambiente.

3) La transmisión de la herencia social de una generación a la siguiente (cultura).

a lo cual WRIGHT (1964) añade:

4) El entretenimiento.

Al parecer toda sociedad posee instituciones que se especializan en estas actividades. Algunas de éstas son el periodismo, la política, la educación, la literatura y las bellas artes. La comunicación de masas puede desempeñar todos esos propósitos simultáneamente de manera más o menos adecuada, afectando así los ordenamientos institucionales dentro de la sociedad. Aunque estas distinciones surgieron en la teoría sociológica y en el análisis político, son también aplicables al examen de fenómenos de organización social más circunscriptos y al de instituciones particulares.

Sin entrar en los detalles de los múltiples (y no siempre bien diseñados) análisis de contenido que han empleado uno u otro tipo de marco institucional, permítaseme efectuar cuatro proposiciones primordiales referidas al análisis de contenido en contextos institucionales.

Las comunicaciones tienden a estar gobernadas por reglas institucionales que prescriben las condiciones en que aquéllas se difunden y utilizan dentro de una organización. Un cheque bancario entregado a un cajero debe ajustarse, en su forma y en su tipo de escritura, a lo que el banco reconoce como cheque válido. Si satisface todas las condiciones requeridas, el cajero está obligado a llevar a cabo la transacción financiera; y conociendo el papel institucional del cheque, un auditor puede inferir cuánto dinero ha sido transferido. De la misma manera, conociendo el papel institucional del diario *Pravda* como órgano oficial del gobierno soviético, puede inferirse que un artículo sin firma sobre cuestiones nacionales soviéticas puede representar la posición de un alto funcionario oficial o del Partido Comunista, e interpretarlo entonces como una transacción oficial que puede tener consecuencias en la política interior. El conocimiento que avala estas inferencias representa los aspectos estables de la red institucional de la cual *Pravda* forma parte. En las investigaciones sobre comunicación de masas, los enfoques institucionales

han puesto el acento sobre las condiciones jurídicas, económicas, sociopolíticas y tecnicoestructurales que conforman los contenidos de los medios, y que a su vez pueden inferirse de las comunicaciones disponibles.

— *Jurídicas*, en el sentido de que los comunicadores deben contar con una licencia para expresarse, pertenecer a una asociación profesional, pagar determinados derechos o, por alguna otra vía, estar autorizados a participar en el proceso de la comunicación. Dentro de las organizaciones sociales, el derecho a utilizar un determinado canal de comunicación es objeto de regulación, y cualesquiera que sean los datos que uno obtenga en dichos contextos, éstos pondrán de relieve qué es lo que la institución considera admisible. Esto incluye tanto las limitaciones legales que ya no son acatadas como las que son aceptadas de forma implícita.

— *Económicas*, en el sentido de que alguien debe pagar los gastos de producción, transmisión y consumo. En los Estados Unidos, los programas de televisión son costeados por la publicidad y todo programa que sale al aire sirve a los intereses de su patrocinador. Los efectos de la propiedad, el monopolio y las conexiones industriales han sido objetivos frecuentes de las inferencias de los especialistas en análisis de contenido. En su mayor parte, pueden evaluarse las comunicaciones en función de sus gastos y beneficios institucionales.

— *Sociopolíticas*, en el sentido de que rara vez las comunicaciones quedan confinadas a ciertos canales institucionales particulares, sino que en general se ven sometidas a las limitaciones derivadas de los códigos éticos formales o de los juicios de valor informales que efectúan los grupos de intereses potencialmente poderosos. El escándalo Watergate es un ejemplo de lo que puede ocurrir cuando se desestiman las poderosas condiciones sociopolíticas de todo acto político. Los periódicos, las emisoras de televisión e incluso las industrias que pretenden perdurar no pueden permitirse el lujo de desagradar a una minoría poderosa o a la elite gobernante en un país. Así pues, en los contextos institucionales, las comunicaciones, en particular las públicas, reflejan las configuraciones de poder predominantes entre los emisores y sus posibles receptores.

— *Técnico-estructurales*, en el sentido de que las comunicaciones deben poder producirse y transmitirse por los medios existentes. La industria cinematográfica y televisiva emplea técnicas de producción en masa enormemente distintas de las que emplean los periódicos o de las que se utilizan en un pequeño circo. Lo que se ve en la pantalla, se lee en el periódico o se contempla en el circo está limitado por sus respectivos medios técnicos de producción y por las formas sociales de organización de esos medios. El análisis de contenido es capaz de inferir algunas de estas condiciones a partir de los mensajes que se puedan recoger.

Creadas y difundidas bajo la égida de las reglas operativas de una institución, *las comunicaciones tienden a reforzar las reglas mediante las cuales han sido creadas* y difundidas. En el interior de los contextos institucionales, la actividad consistente en enunciar algo suele ser más importante que lo que se enuncia. Tal vez vengan a la mente del lector los discursos pronunciados en ocasión de ciertas ceremonias: su propósito primordial es trasladar una actuación ritual de un escenario al siguiente, demostrando así, con la realización de la secuencia, el éxito de una institución.

Las *propiedades de un medio* en cuanto al registro y difusión de la información *ejercen un profundo efecto sobre la naturaleza de las instituciones que pueden sustentarse mediante las comunicaciones a través de ese medio.* Comparando las comunicaciones orales con las escritas, INNIS (1951) llegó a la conclusión de que la escritura tiene por efecto cristalizar las tradiciones, y demuestra ser más permanente y fiable, por lo cual es capaz de sustentar imperios que extienden su control a enormes zonas geográficas, mucho mayores que lo que permitirían las comunicaciones orales. La radio y la televisión, con su transmisión prácticamente instántanea a gran distancia, tienden a apoyar el desarrollo de organizaciones dispersas en una amplia zona geográfica, pero también son mucho menos capaces de conservar el acceso a la historia. Los medios de comunicación, o algunas de sus propiedades, se consideran, pues, como el principal agente de cambio social y de desarrollo de las estructuras sociales. En el proceso de mediación el control se confiere a determinadas formas institucionales. El análisis de contenido de estos medios en el interior de contextos institucionales puede dar origen a inferencias relacionadas con la competencia de las diversas modalidades de

comunicación en lo que se refiere a alcanzar el predominio, así como con el tipo de cambios sociales que se aceleran o retrasan, o con la forma en que se distribuye el poder en una sociedad.

Transmitidas a través de canales institucionales, *las comunicaciones tienden a adoptar la sintaxis y la forma que dichos canales pueden transmitir con mayor eficiencia.* Los análisis de contenido han arrojado luz sobre los cambios sistemáticos que se producen en el contenido cuando se lleva un libro a la pantalla (ASHEIM, 1950); o lo que ocurre con las noticias y espectáculos de contenido discutible, los llamados "mecanismos de discriminación y censura" [*gate-keeping*] (WHITE, 1964); o la forma en que la noticia se fabrica y no se transmite (GIEBER, 1964); o el papel social de las modelos que posan para las portadas de revistas, como función de los canales de distribución (GERBNER, 1959); o la manera en que las expectativas acerca de una institución pueden conformar las peticiones de los usuarios (KATZ y otros, 1967). En su mayoría, los estudios sobre comunicación de masas que emplean esta perspectiva avalan la opinión de que la necesidad de repetición (producción en masa) tiende a preservar y fortalecer los estereotipos sociales, los prejuicios y las ideologías, en lugar de modificarlos (ADORNO, 1960). No obstante, han existido propuestas para anticipar los grandes cambios sociales a partir del análisis de contenido de la literatura de vanguardia, centrándose en las comunicaciones que pueden socavar o eludir los controles institucionales (BELL, 1964).

4. La lógica del proyecto del análisis de contenido

A partir de aquí se hablará de la naturaleza secuencial de los diseños de análisis de contenido: la conformación de los datos, su reducción, la inferencia y el análisis llevan al especialista en análisis de contenido a procesos de validación directa, de verificación de la correspondencia con otros métodos y de verificación de sus hipótesis.

PROCESAMIENTO DE LA INFORMACION CIENTIFICA

Probablemente, la actitud más comprometida y clara que debe adoptar el analista de contenido tiene que ver con su manera de procesar la información. Los descubrimientos científicos nunca se aceptan al pie de la letra, ya tengan como finalidad la mejor comprensión de un fenómeno o el perfeccionamiento de determinadas condiciones de vida. El investigador es el responsable del proceso que conduce a dichos descubrimientos: debe describir las condiciones en que obtiene los datos, justificar los pasos analíticos seguidos y procurar en todo momento que el proceso no sea tendencioso, en el sentido de favorecer cierta clase de hallazgos en detrimento de otros. Es necesario que el proceso sea explícito para que otros puedan evaluar su labor, reproducir dicho proceso o restringir sus hallazgos.

La red de pasos analíticos mediante los cuales se procesa la

información científica se denomina "proyecto o diseño de la investigación". Este diseño da cuenta de la manera en que se obtienen los datos y de lo que se hace con ellos en el curso del análisis, y proporciona instrucciones a otras personas acerca de lo que deben hacer si pretenden reproducir los resultados. *Todo informe de investigación debe contener una descripción del proyecto de investigación.*

En el caso del análisis de contenido, más aún que en el de otras técnicas, el proyecto o diseño de investigación en su conjunto debe adecuarse al contexto del cual provienen los datos o con respecto al cual se analizan éstos. Por ejemplo, si se trata de inferir la psicopatología de un paciente neuropsiquiátrico a partir de sus respuestas a determinadas preguntas, no parece que tenga sentido dividir esas respuestas en palabras sueltas, mezclar éstas y extraer una muestra al azar para su análisis. Este procedimiento sólo se justificaría si la información deseada estuviera contenida en la aparición de palabras aisladas. Lo más probable es que deba tomarse como unidad el par pregunta-respuesta, por poseer una organización interna que se estima significativa. Otro caso: si se intenta crear el perfil sociológico de un periódico, de nada valdría aplicar a su contenido categorías del tipo de las que aparecen en un test de inteligencia. Las categorías tienen que justificarse en función de lo que se conoce en el contexto de los datos. Los proyectos de investigación para el análisis de contenido tienen que ser *sensibles al contexto.* Debe existir alguna correspondencia, explícita o implícita, entre el procedimiento analítico y las propiedades pertinentes del contexto.

Los proyectos de investigación para el análisis de contenido tienden a ser de naturaleza *secuencial.* En ellos, cada paso sigue a uno anterior, y las decisiones sobre un procedimiento determinado no se toman (ni se consideran) según el resultado del procedimiento siguiente. De esto se desprende que cualquier error incorporado al proyecto y que no haya sido detectado permanecerá en él hasta el fin. En los procesamientos secuenciales de la información los errores son, por lo tanto, acumulativos o multiplicativos. Si uno de los componentes de un proyecto de investigación es defectuoso (por ejemplo, porque destruye información pertinente) ningún componente posterior compensará esta pérdida ni permitirá recuperar lo destruido. Si los observadores no registran sus observaciones de manera fiable, por muy cuidadosos que sean los análisis posteriores darán resultados que no resultarán fiables.

TIPOS DE PROYECTOS

Si contemplamos los proyectos de investigación para el análisis de contenido de una manera global, podemos distinguir tres tipos. La tipología depende en gran medida de la manera en que los resultados de un análisis de contenido se insertan en trabajos de investigación más amplios.

Los más básicos son los *proyectos para evaluar* ciertos fenómenos en el contexto de los datos existentes. Estos responden de forma directa a la definición de análisis de contenido, y se utilizan cuando *el análisis de contenido es el único método* utilizado. Podrían distinguirse, a la vez, aquellos casos en que se evalúa algún parámetro o se pone a prueba alguna hipótesis entre diversos parámetros estimados; no obstante, la forma en que estas dos clases de proyectos de investigación se relacionan con la realidad (concretamente, el hecho de que los hallazgos empíricos sean interpretados como indicativos del contexto) los convierte en

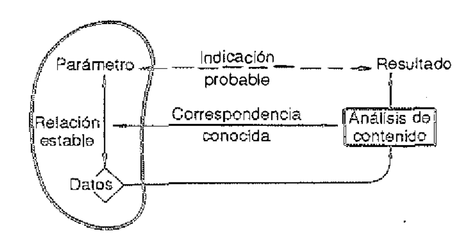


Figura 3. Diseño de análisis de contenido para estimación.

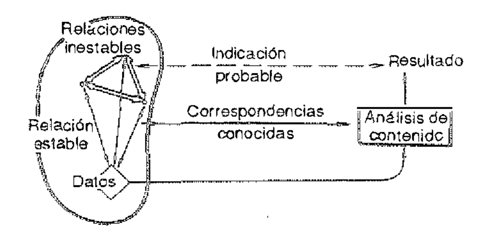


Figura 4. Diseño de análisis de contenido para inferir relaciones.

esencialmente iguales. Ambas situaciones se exponen gráficamente en las figuras 3 y 4.

Entre los ejemplos de las estimaciones de parámetro único podemos mencionar las inferencias sobre el nivel de angustia de un paciente psiquiátrico en el transcurso de una entrevista, la evaluación de las actitudes de un orador o de posiciones o valores ideológicos, el intento de calibrar el espíritu bélico de la población de un país mediante las emisiones radiofónicas programadas por su círculo dirigente o las estimaciones de la cantidad de atención pública dedicada a un problema social. Entre los ejemplos de hipótesis incluidas en el contenido, podemos mencionar las correlaciones entre: 1) las características de la personalidad de personajes de la televisión indicativos de estereotipos y fórmulas, y 2) los estudios sobre el predominio y el desplazamiento de determinados temas de interés a lo largo del tiempo.

También podría diferenciarse entre: 1) aquellos casos en que el analista del contenido tiene motivos para creer que el proyecto de investigación está relacionado con la vinculación entre los datos y su contexto, o constituye un modelo de dicha relación, o la reproduce, y por ello acepta los resultados del proceso como representativos de los fenómenos contextuales, y 2) los casos en que el éxito que ha logrado en el pasado con el método lo lleva a confiar en lo que representan sus resultados, independientemente de la correspondencia estructural que exista entre su proyecto de investigación y las relaciones contextuales que abarcan los datos. Estas dos opciones no son mutuamente excluyentes. Los especialistas en análisis de contenido tienden a rechazar los resultados que les parecen inadmisibles o los procedimientos de investigación que abarcan procesos evidentemente ajenos, a su juicio, a lo que saben sobre la fuente de los datos. A menudo, el analista conoce una pequeña porción de su objetivo y una pequeña porción de las relaciones.

En el interior de este diseño es decisivo que el analista utilice todo el conocimiento de que dispone sobre el sistema que le interesa para interpretar *un solo* conjunto de datos no estructurados o simbólicos. No se apoya en otros métodos para dar validez a sus resultados ni puede considerar simultáneamente diferentes conjuntos de datos, en relación con los cuales podría obtener conocimientos adicionales.

Un segundo tipo son los *proyectos para poner a prueba la posibilidad de sustituir* un método a través de un análisis de con-

tenido. En este caso, se aplican dos o más métodos a los mismos datos o a datos diferentes obtenidos a partir de la misma situación, con el fin de verificar si ambos proporcionan resultados comparables, o en el caso de que haya en juego más de dos métodos, cuál es el mejor. La figura 5 constituye la representación gráfica de este diseño.

Entre los ejemplos podemos mencionar las comparaciones del cociente de alivio de la incomodidad aplicadas a datos de entrevistas y el juicio de un equipo de psiquiatras sobre el nivel de angustia de un sujeto (DOLLARD y MOWRER, 1947), o los intentos de establecer correlaciones entre diversos métodos de medición de la atención pública. Por lo general, uno de los dos métodos se estima más válido que el análisis de contenido, pero este último es en algún aspecto más ventajoso, ya sea por su precio, por la rapidez con que se puede realizar o por su negativa a la intromisión. Una correspondencia perfecta tal vez no nos diga mucho acerca de la validez de uno u otro método, pero sí puede indicar una equivalencia funcional, y por lo tanto su intercambiabilidad o la posibilidad de sustituirlos mutuamente (véase el apartado sobre la validez). Gran parte de la investigación del análisis de contenido intenta realizar la búsqueda de técnicas para inferir, a partir de datos simbólicos, lo que mediante el uso de otras técnicas resultaría imposible conocer, o sería excesivamente costoso, o provocaría un alto grado de intromisión.

Con frecuencia se aplica toda una batería de métodos distintos, incluido el análisis de contenido, a datos derivados de la misma situación. Sin saber *a priori* cuál de ellos es "mejor" o más válido, las posibles altas correlaciones entre sus resultados se interpretan como un indicio de que evalúan los mismos fenómenos subyacentes. Este diseño de investigación sirve de base a lo que se ha dado en llamar *operacionalismo múltiple* (WEBB y otros, 1966). Exige que los resultados de la investigación sean coherentes en los diferentes métodos empleados, todos los cuales se pretenden sensibles a los mismos fenómenos. Procura, pues, impedir que los descubrimientos sean simplemente el resultado artificial de un único método, quizá desvinculados por completo de los resultados de otras investigaciones.

Los *proyectos para poner a prueba hipótesis*, que aparecen en la figura 6, comparan los resultados de un análisis de contenido con datos obtenidos de manera independiente y con datos sobre fenómenos no inferidos mediante la técnica. Muy a menudo, los

análisis de contenido son sólo una parte de un esfuerzo de investigación más amplio. Pueden existir varias clases de datos disponibles, y puede que sólo algunos de ellos no estén estructurados o vinculados con comunicaciones simbólicas a las que deba aplicarse un análisis de contenido. Este proyecto de investigación permite comprender las relaciones que pueden existir entre los fenómenos de los que se ocupa un análisis de contenido y sus condiciones circundantes. Por ejemplo, pueden someterse a un análisis de contenido los programas de televisión para averiguar la presencia de ciertas características, y luego cotejar esto con datos obtenidos en entrevistas sobre las preferencias del público en materia de programas, con el fin de cerciorarse si existe una relación entre la estratificación social y la característica seleccionada. La presencia de correlaciones entre las intenciones declaradas y el comportamiento real de un sujeto pueden poner de relieve si el contenido cumple propósitos meramente estratégicos. Las correlaciones entre el contenido de un mensaje, tal como se deduce del análisis de contenido, con una variedad de índices de comportamiento también proporciona conocimientos acerca de las condiciones antecedentes y de los efectos. Así, la cantidad de hijos de las familias que protagonizan las obras de ficción de las revistas populares se correlacionó con los índices de natalidad (MIDDLETON, 1960); la cobertura periodística de determinados temas sintetizados en cifras, como los índices de precios al consumidor, las tasas de desempleo y los índices ecológicos, se analizaron con respecto a las fluctuaciones de la opinión pública y de las cifras reales (ZUCKER, 1978). Asimismo, las correlaciones entre la frecuencia de relatos de amplia divulgación sobre suicidios y las catástrofes aéreas revelaron la existencia de una relación causal o "desencadenante" (PHILLIPS, 1978).

COMPONENTES DEL ANALISIS DE CONTENIDO

Examinando ahora con detalle el proyecto de investigación para el análisis de contenido, cabe distinguir varios componentes o pasos diferentes en este proceso:

- Formulación de los datos.
  - Determinación de las unidades.
  - Muestreo.
  - Registro.
- Reducción de los datos.
- Inferencia.
- Análisis.

Para lo cual se recurre a :

- Validación directa.
- Verificación de la correspondencia con otros métodos.
- Verificación de las hipótesis respecto de otros datos.

La figura 7 muestra gráficamente cómo pueden ensamblarse estos componentes en un diseño de investigación de análisis de contenido.

En los posteriores capítulos describiremos con cierto detenimiento estos componentes a la manera de un equipo de herramientas, incluyendo los problemas que presentan y las soluciones prácticas, de manera que todo análisis de contenido pueda adaptarlos a sus necesidades. En el resto de este capítulo realizaremos algunos comentarios generales adicionales sobre estos componentes.

Elaboración de los datos

Un dato es una unidad de información registrada en un medio duradero, que se distingue de otros datos, puede analizarse mediante técnicas explícitas y es pertinente con respecto a un problema determinado. Así definidos, los datos no son "hechos" absolutos: reciben una forma particular con una finalidad particular, y gran parte de los trabajos del análisis de contenido están destinados a dar forma analizable a una información no estructurada y vicaria, con frecuencia contingente.

Los datos deben transportar información, en el sentido de suministrar el nexo entre las fuentes de información y las formas simbólicas espontáneas, por un lado, y las teorías, modelos y conocimientos concernientes a su contexto, por el otro. Los datos deben ser representativos de fenómenos reales.

La necesidad de registrar los datos en un medio duradero se

desprende del requisito de la reproducibilidad. Sólo los registros con cierto grado de perdurabilidad son reanalizables. El sonido de la voz humana se desvanece poco después de emitida y no puede, por lo tanto, considerarse como un dato. Con el fin de someterla a un análisis de contenido, el habla humana debe aparecer escrita o, al menos, registrada en cinta magnetofónica.

La necesidad de que los datos sean analizables mediante técnicas explícitas los relaciona con la capacidad analítica del investigador. Pocos años atrás, la palabra escrita no podía considerarse un dato, ya que requería intérpretes humanos para conferirle sentido. Con el advenimiento de los procesadores de datos lingüísticos, tanto las palabras sueltas, como las oraciones, párrafos, capítulos y libros enteros son aceptados como datos. Una tendencia permanente de la historia de la ciencia es someter un número creciente de fenómenos a la medición y el análisis.

En el análisis de contenido, los datos emergen por lo general a partir de formas simbólicas complejas, enunciadas en un lenguaje espontáneo. Las historietas, apuntes privados o diarios íntimos, obras literarias y teatrales, telenovelas, anuncios publicitarios, películas cinematográficas, discursos políticos, documentos históricos, interacciones en pequeños grupos, entrevistas o acontecimientos sonoros, tienen cada cual su propia sintaxis y semántica, y rara vez es posible analizar estos fenómenos en su manifestación original. En el interior de estas formas no estructuradas deben cumplirse los siguientes requisitos:

- Los fenómenos de interés deben distinguirse y dividirse en unidades de análisis separadas, lo cual plantea el problema de la *determinación de las unidades*.
- Estas unidades pueden presentarse en una cantidad tan grande que no permita un manejo fácil, lo cual plantea el problema del *muestreo* de una porción más pequeña a partir de todas las unidades posibles.
- Cada unidad debe codificarse y describirse en formas analizables, lo cual plantea problemas de *registro*.

La determinación de unidades, el muestreo y el registro están de alguna manera interconectados. Para muestrear una proporción de una clase de unidad y otra proporción de otra clase, por ejemplo, se necesita distinguir entre las dos clases, lo cual constituye un aspecto importante del proceso de registro. La determinación

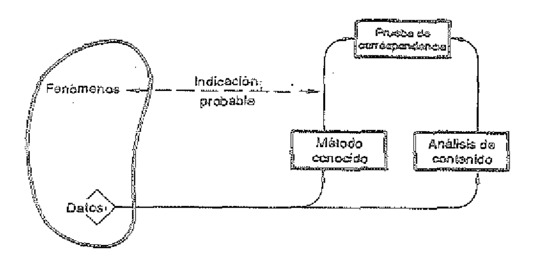


Figura 5. Diseño de análisis de contenido para comparar métodos diferentes.

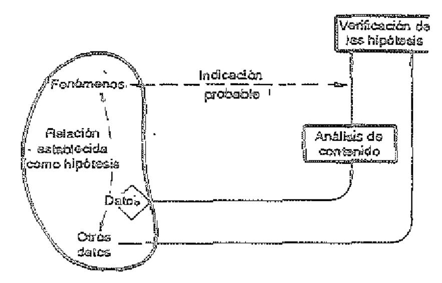


Figura 6. Diseño de análisis de contenido para verificar hipótesis.

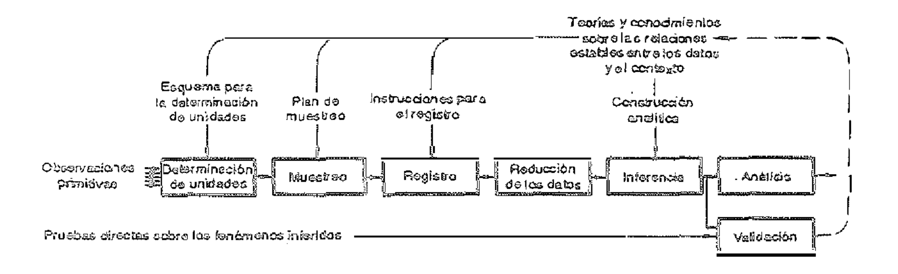


Figura 7. Procedimientos utilizados en el análisis de contenido.

de unidades también puede llevarse a cabo durante la fase de registro del análisis de contenido. Ejemplo de esto sería el caso en que se distinguen las fases de un proceso grupal a raíz de que entre una y otra fase las diferencias de comportamiento son mayores que las que existen en cada fase. Cuando se analiza todo el universo de datos, no es necesario el muestreo.

Reducción de los datos

No nos extenderemos acerca de la reducción de los datos en este capítulo porque a veces no presenta ningún problema, o bien, otras veces, puede tratarse dentro del tema del análisis, del que nos ocuparemos más adelante. Aunque en cualquier punto del proyecto de investigación puede existir reducción de los datos, ésta se relaciona principalmente con la facilidad de los cálculos, y procura adecuar la forma de los datos disponibles a la exigida por la técnica analítica. La reducción puede ser estadística, algebraica, o tener como única finalidad la omisión de los detalles irrelevantes.

Inferencia

La inferencia es, desde luego, la razón de ser de todo análisis de contenido y la describiremos luego con minuciosidad. "Abarca" todo el saber que debe poseer el analista de contenido acerca del modo en que los datos se relacionan con su contexto, saber que se verá fortalecido por el éxito de cada inferencia. En las figuras, esto se indica mediante la flecha que señala el resultado de los trabajos de validación.

Análisis

El análisis se ocupa de los procesos más convencionales de identificación y representación de las pautas más notables, estadísticamente significativas o que por algún otro motivo dan cuenta de los resultados del análisis de contenido o los describen. Por definición, al análisis no se le exige que sea sensible al contexto, como ocurre con los procedimientos anteriores.

Según esto, un diseño de investigación debe ser reproducible, y lo que debe ser válido para la totalidad, debe serlo también para cada uno de los detalles. Cuando se estandariza un componente (como el análisis de contingencia mediante programas de ordenador, que no puede someterse a fluctuaciones aleatorias), queda prácticamente asegurada la reproducibilidad. Pero si los procedimientos analíticos implican a individuos, invariablemente se introducirán errores e incertidumbres. La evaluación de estos erro-

res se examinará cuando nos ocupemos de la fiabilidad; no obstante, digamos que la condición primordial para la reproducibilidad de un proyecto de investigación es que cada uno de sus componentes sea descrito de forma explícita.

5. Determinación de las unidades

En este capítulo se explican las unidades empleadas en el análisis de contenido, su finalidad y los procesos necesarios para definirlas —ya se trate de unidades de muestreo, de registro o de contexto— de un modo eficiente y fiable.

La primera tarea de toda investigación empírica consiste en decidir qué se ha de observar y registrar, y lo que a partir de ese momento será considerado un dato. Hay buenas razones para utilizar el plural "datos" en lugar del singular, ya que toda investigación empírica abarca una multitud de unidades portadoras de información. La determinación de las unidades comprende su definición, su separación teniendo en cuenta sus respectivos límites y su identificación para el subsiguiente análisis.

TIPOS DE UNIDADES DE ANALISIS

Relacionar el campo de las observaciones y el de los mensajes plantea muchos problemas epistemológicos que no podemos abordar aquí, salvo para decir que las unidades nunca son absolutas: surgen de la interacción entre la realidad y su observador; son una

función de los hechos empíricos, de las finalidades de la investigación y de las exigencias que plantean las técnicas disponibles. En el análisis de contenido merecen distinguirse tres clases de unidades: *unidades de muestreo, unidades de registro* y *unidades de contexto*. Nos detendremos en sus propósitos analíticos y su empleo, tras lo cual examinaremos cinco formas distintas de definir estas unidades.

Unidades de muestreo

Las unidades de muestreo son aquellas porciones de la realidad observada, o de la secuencia de expresiones de la lengua fuente, que se consideran independientes unas de otras. Aquí, "independientes" es sinónimo de no relacionadas, no ligadas entre sí, no ordenadas o libres, de modo que la inclusión o exclusión de una unidad de muestreo cualquiera como dato en un análisis carece de consecuencias lógicas o empíricas en lo que se refiere a las elecciones entre otras unidades. Una definición estadística de dichas unidades subrayaría que "en su interior hay muy poca libertad para la variación, pero hay mucha libertad para la variación en sus límites" (POOL, 1959, pág. 203). No obstante, la elección de unidades de muestreo rara vez está motivada por fines estadísticos.

Puede servirnos de ejemplo un discurso político. Si el oyente común reacciona frente a él de una manera holística o total, diciendo que el orador le gusta o le disgusta, que atrae o no su simpatía, para el analista político, en cambio, el discurso aborda varias cuestiones separadas. Puede entonces dividirlo y, haciendo caso omiso de las conexiones entre esas partes, sondear los detalles estructurales de algunas de ellas o de todas con el fin de poner de relieve las actitudes o las pautas de razonamiento del orador. Un lingüista, por el contrario, probablemente dividirá el discurso en oraciones. Como no hay reglas gramaticales que consideren una construcción oracional dependiente de otra, no verá la necesidad de considerar unidades mayores que la oración. Para él, el conjunto de oraciones contiene toda la información pertinente para explorar la actuación lingüística del orador. Por otra parte, un programa de ordenador que se aplique al cómputo de palabras, al dejar de lado su posición sintáctica, determinará las unidades del texto de un modo que resultará inútil para el lingüista. Análogamente, el analista político considerará de escaso interés el conjunto de oraciones establecidas por el lingüista, ya que sus conoci-

mientos proceden de la organización semántica y retórica del discurso, que es ignorada por el lingüista al determinar sus unidades.

Como es obvio, las unidades de muestreo son fundamentales para realizar un muestreo, ya que éste se extrae, unidad por unidad, del universo de unidades muestrales. Las unidades muestrales son importantes también para la estadística inferencial. Los objetos que deben computarse tienen que ser independientes entre sí, pues de lo contrario sus frecuencias carecerían de significado. Los investigadores de encuestas que hacen amplio uso de la estadística inferencial se empeñan en asegurarse de que sus entrevistadores no interactúen entre sí. Los experimentadores se cercioran, a su vez, de que los sucesos que manipulan para poner a prueba las hipótesis estadísticas sean independientes.

Las interdependencias entre las unidades de muestreo, si es que existen, no sólo se pierden en el muestreo, sino que además confunden los hallazgos. Supóngase que durante un estudio de la violencia televisiva se aprecia en un programa que un incidente violento desencadena una larga serie de enfrentamientos violentos, que constituye el tema del drama, mientras que en otro programa hay muchos enfrentamientos violentos provocados de forma aislada y que no son consecuencia de ningún otro, girando el tema en torno a alguna otra cosa. ¿De qué manera podrían definirse y enumerarse unidades de muestreo significativas para analizar la violencia? ¿Debe considerarse que cada serie de enfrentamientos violentos es una única unidad de muestreo, o debe dividirse en varias? ¿Se otorgará a cada una de ellas igual peso? Quizá los enfrentamientos violentos no constituyan unidades significativas de muestreo. A menos que se encuentren unidades independientes en lo referente al fenómeno que interesa, si se divide un mensaje de alto grado de organización en unidades muestrales separadas, invariablemente se deformará la información contenida en los datos resultantes.

Unidades de registro

Las unidades de registro se describen por separado, y pueden considerarse partes de una unidad de muestreo que es posible analizar de forma aislada. Aunque las unidades de muestreo tienden a poseer límites físicamente discernibles, las distinciones entre las unidades de registro, en cambio, son el resultado de un trabajo descriptivo. HOLSTI (1969, pág. 116) define una unidad de registro

como "el segmento específico de contenido que se caracteriza al situarlo en una categoría determinada". Las dependencias que podrían existir dentro de las unidades de muestreo se mantienen en la descripción individual de su unidad de registro.

Supóngase que se quiere extraer una muestra de los programas televisivos de ficción, con el objeto de estudiar la población de los personajes televisivos. Estos programas tienen un comienzo y un final bien definidos, y por ello constituyen unidades de muestreo naturales, respecto de las cuales es posible decidir fácilmente su inclusión en la muestra o su exclusión, sin tener en cuenta el contenido de dichos programas. Los personajes distan de ser independientes uno de otro: interactúan, cada uno de ellos se define con relación al otro, y asumen roles específicos dentro de un guión que posee un alto grado de organización. A raíz de estas interdependencias, sería imposible dividir un programa en segmentos que tuvieran significado individual, uno por cada personaje; lo que sí es posible es describir a cada personaje individualmente, incluyendo la descripción de la posición que asume dentro de la red de relaciones interpersonales, su orden de aparición y las interacciones en que participa. En el curso de esta descripción, los personajes individuales se convierten en unidades de registro que pueden analizarse de forma individual y colectiva, aunque no siempre de forma significativa.

Otro motivo para elegir unidades de registro que difieran de las unidades de muestreo es que estas últimas con frecuencia son demasiado amplias, ricas o complejas como para utilizarlas en la descripción. Por ejemplo, en el caso de una película cinematográfica que hace un uso creativo de material documental, es difícil clasificar dicho material como real o ficticio: el film contiene ambas condiciones. No obstante, describiendo unidades menores (escenas, planos o imágenes individuales, por ejemplo), pueden obtenerse unidades de registro codificables de manera inequívoca.

Es posible, aunque a menudo difícil, describir las unidades de registro de tal forma que pueda reconstruirse toda la unidad de muestreo de la que forman parte. Para ello se necesita que en el conjunto de unidades de registro se mantenga la información pertinente acerca de la organización de la unidad de muestreo. El analista cuenta entonces con la opción de analizar diferentes niveles de unidades.

Una costumbre habitual en el análisis de contenido es definir estructuralmente unidades más amplias. En dichas definiciones,

debe ser posible identificar una cierta estructura entre las unidades menores para que la unidad más amplia sea admisible con vistas al análisis. Un buen ejemplo es el esquema llamado "acción-actor-objetivo", donde cada uno de éstos puede caracterizarse de forma individual y, por lo tanto, analizarse separadamente, pero los tres deben co-ocurrir en una unidad de registro si la meta es analizar las co-ocurrencias que se producen entre las unidades.

Unidades de contexto

Las unidades de contexto fijan límites a la información contextual que puede incorporarse a la descripción de una unidad de registro. Demarcan aquella porción del material simbólico que debe examinarse para caracterizar la unidad de registro. Definiendo una unidad de contexto más amplia para cada unidad de registro, el investigador reconoce y explica el hecho de que los símbolos codeterminan su interpretación y de que extrae sus significados, en parte, del medio inmediato en el que se presentan. Las unidades de contexto no necesitan ser independientes ni descriptibles de forma aislada; pueden superponerse, y contienen numerosas unidades de registro.

Una investigación llevada a cabo por GELLER y otros (1942) demostró hasta qué punto la caracterización de una unidad de registro, y en última instancia los resultados de toda una investigación, dependen de la magnitud de la unidad de contexto. En dicho estudio, los autores solicitaron a los sujetos que juzgaran de qué manera se evaluaban símbolos del tipo "democracia", utilizando como unidades de contexto una oración, un párrafo, tres oraciones y un artículo completo. Aunque en general los cuatro métodos coincidieron entre sí en cuanto a la dirección de la orientación (favorable, neutral, desfavorable), difirieron en cuanto a su intensidad. A medida que aumentaba el tamaño del contexto, disminuía significativamente el número de evaluaciones neutrales. Evidentemente, el contexto de un símbolo contiene gran cantidad de información evaluativa.

No obstante, problemas de fiabilidad y de eficiencia impiden que se recurra a contextos más amplios. Si se quiere describir cómo se tratan los personajes en el contexto de una novela, debe leerse primero el libro completo y luego situar a cada personaje en la categoría correspondiente. Este proceso no sólo exige tiempo sino que a menudo resulta poco fiable, ya que diversos sujetos

pueden abordar la novela de diferentes maneras, y para emitir un juicio es preciso tener en cuenta la novela en su totalidad. Más eficiente y fiable, aunque menos significativo, sería recorrer un documento oración por oración, o bien, si se trata de otros datos, recorrer las sucesivas escenas, diálogos o planos.

En otros textos sobre el análisis de contenido (por ejemplo, HOLSTI, 1969) se mencionan también *unidades de enumeración*. La importancia de estas unidades estaba relacionada con la definición primitiva del análisis de contenido, que requería que éste fuese cuantitativo (BERELSON, 1952), lo cual significaba simplemente que debía darse cuenta de los datos numéricamente, en función de la frecuencia de su aparición, del espacio (centímetros de columnas en las noticias de periódicos), del tiempo (minutos de emisiones radiofónicas) o, en fin, de las características tipográficas (tipo de letra usado en los títulos o tamaño de las figuras o fotos). Sin embargo, el estatuto conceptual de estas unidades es confuso. En el análisis de contenido, las cantidades pueden tener hasta tres orígenes diferentes. En primer lugar, pueden ser el resultado del cómputo de las apariciones reiteradas. Esto asocia *una magnitud con una clase de unidades de registro idénticas* y lo único que hace es reducir el esfuerzo analítico. En segundo lugar, medidas como el tamaño de una fotografía o la cantidad de centímetros de columna que abarca un artículo pasan a ser *descriptivas de una sola unidad de registro*. En tercer lugar, la cantidad de ejemplares impresos de un periódico o clasificaciones de programas televisivos como la de Nielson, se asocian con toda una unidad de muestreo y *son compartidas por todas las unidades de registro* que componen dicha unidad de muestreo. En los antiguos análisis de contenido, el codificador debía cuantificarlo todo, y la distinción entre estas tres clases de cantidades se hacía confusa. Las cantidades del primer tipo salen a relucir en el proceso del análisis de los datos. Por ejemplo, al obtener coeficientes de correlación y medidas de distancia, las frecuencias y magnitudes de las clases de unidades son meramente computacionales: el codificador no necesita preocuparse en absoluto de ellas. Las cantidades del segundo tipo sí preocupan a los codificadores, por cuanto deben medirse en cada unidad de registro. No obstante, los centímetros cuadrados o lineales, o la cantidad de líneas, no exigen ningún concepto especial: tienen el carácter de una categoría numérica, del mismo modo que los valores de un diferencial semántico o de una escala de intensidad. Las cantidades del tercer tipo tienden a suministrar-

las otro tipo de fuentes y rara vez proceden de los juicios del propio codificador. Así, pues, si bien es importante saber qué es lo que se cuantifica en el análisis de contenido, no es preciso que nos detengamos aquí particularmente en las unidades de numeración.

Por consiguiente, las unidades se distinguen de acuerdo con la función que desempeñan en el análisis de contenido. Las unidades de muestreo interesan para el muestreo y sirven de base para los estudios de tipo estadístico. Las unidades de registro, en su conjunto, son portadoras de la información dentro de las unidades de muestreo y sirven de base para el análisis. Y las unidades de contexto se refieren al proceso de descripción de las unidades de registro.

PROCEDIMIENTOS PARA DEFINIR LAS UNIDADES

Pese a las diferencias funcionales que existen entre ellos, la mayoría de los análisis de contenido aplican uno o más de cinco procedimientos distintos para fijar e identificar estas unidades:

- Unidades físicas.
- Unidades sintácticas.
- Unidades referenciales.
- Unidades proposicionales (y núcleos de significado).
- Unidades temáticas.

A continuación pasaremos a describir cada una de éstas.

Unidades físicas

Algunas unidades parecen tan obvias que casi no vale la pena dedicarles atención especial: un libro, un informe financiero, un tema tratado por un periódico, una carta, un poema o un *poster*, son todas ellas unidades físicamente determinadas, y si parecen obvias es porque el límite del mensaje que contienen coincide con el límite del medio por el cual se transmite.

Aun cuando los sucesos sean continuos o el flujo de las expresiones de la lengua fuente muestre pocos límites naturales, pueden imponérseles divisiones físicas. Las unidades físicas dividen un medio de acuerdo con el tiempo, la longitud, el tamaño o el volumen, y no de acuerdo con la información que transmiten. OSGOOD

(1959) tomó muestras de las páginas del diario íntimo de Goebbels. EKMAN y FRIESEN (1968) utilizaron como mínima unidad de registro las imágenes de películas cinematográficas. DALE (1937) analizó los noticiarios cinematográficos tomando como unidad el pie (30 cm aproximadamente) de película y ALBIG (1938) suministró a un grupo de observadores un reloj y les pidió que sintetizaran cada minuto de emisión radial. Las unidades temporales son corrientes en los estudios de la conducta interpersonal (WEIK, 1968). Se consigue el mismo efecto fijando una cuadrícula sobre una fotografía y describiendo a continuación cada uno de los cuadros así formados.

Unidades sintácticas

Las unidades y elementos sintácticos son "naturales" en relación con la gramática de un determinado medio de comunicación. No exigen emitir juicios sobre el significado.

La palabra es la unidad más pequeña de los documentos escritos, y en lo que concierne a la fiabilidad, la más segura. Muchas son las investigaciones que se han basado en las palabras o símbolos como unidades: el estudio mundial sobre la atención de Lasswell (LASSWELL, 1941; LASSWELL y otros, 1952), numerosos intentos de localización de textos literarios (YULE, 1944; MOSTELLER y WALLACE, 1964), el análisis del estilo (MILES, 1951; HERDAN, 1960), las inferencias psicodiagnósticas (DOLLARD y MOWRER, 1947), y las investigaciones sobre la legibilidad (FLESH, 1948-1951; TAYLOR, 1953) son ejemplos de ello.

Unidades sintácticas de medios no verbales son los espectáculos de televisión (tal como se los enumera en la *TV Guide* o guía de la televisión norteamericana), los diálogos en las obras teatrales, cada una de las noticias que se transmiten en una emisión radiofónica o los planos en las películas. Reconocer las unidades sintácticas requiere familiaridad con el medio. Estas unidades son más naturales que las físicas porque utilizan distinciones establecidas por la propia fuente.

Unidades referenciales

Las unidades pueden definirse a partir de determinados objetos, sucesos, personas, actos, países o ideas a los que se refiere una expresión. De esta manera, el 37º presidente de los Estados Uni-

dos puede ser aludido llamándolo simplemente "él" (cuando el contexto establece inequívocamente de quién se trata), o "el primer presidente norteamericano que visitó la China", o "Richard M. Nixon", o "Dick el tramposo" (el apelativo popular que usaban sus detractores) o "el ocupante de la Casa Blanca entre 1969 y 1974". Todas estas expresiones designan a la misma persona, aunque de manera distinta, y poco importa que la referencia abarque una palabra o muchas, que sea directa o indirecta.

Las unidades referenciales son indispensables cuando se trata de cerciorarse del modo en que se describe un fenómeno existente. Gran parte de los primeros trabajos sobre análisis de símbolos (POOL, 1959) definían éstos (que habitualmente eran palabras sueltas) por sus valores en relación con lo denotado y sus valores explorados, sus atributos y los calificativos asociados con ellos. Si se pretende inferir las actitudes, preferencias y creencias de los autores, se requieren designaciones referenciales de los objetos actitudinales de interés, como sucede en todo trabajo que establezca perfiles de clases particulares de individuos (los héroes, los maestros, los hispanoamericanos).

Unidades proposicionales (y núcleos de significado)

El uso exclusivo de unidades referenciales implica que el lenguaje de datos reconoce simplemente los objetos y sus atributos: no aborda todas las complejidades de la lengua natural. Una manera de establecer unidades algo más complejas es exigir que posean determinada estructura. Por ejemplo, OSGOOD y otros (1956) sugirieron estudiar todas las proposiciones que puedan enunciarse en una de las dos formas siguientes:

Objeto actitudinal / conector verbal / término de significado corriente.
Objeto actitudinal$_1$ / conector verbal/ objeto actitudinal$_2$.

De acuerdo con esto, la oración: "El estaba librando una batalla perdida contra el orden establecido", se transformaría así:

El estudiante-dirigente / estaba librando una batalla / contra el orden establecido.
El orden establecido / es / poderoso.
El estudiante-dirigente / pierde / la batalla.

Esta forma de establecer los núcleos de significado de una oración compleja dividiéndola en unidades proposicionales es la base para evaluar el análisis de aseveraciones.

Holsti (en NORTH y otros, 1963, pág. 137) instruyó a los codificadores para que editaran y reformularan los documentos políticos teniendo en cuenta un marco de referencia de la acción que contenía las siguientes unidades:

- El sujeto que percibe y los modificadores incorporados.
- El sujeto que percibe, distinto del autor del documento, y los modificadores incorporados.
- Lo percibido y los modificadores incorporados.
- La acción y los modificadores incorporados.
- El objeto sobre el cual se actúa (distinto del actor-objetivo) y el modificador incorporado.
- El verbo auxiliar modificador.
- El objetivo y los modificadores incorporados.

Cada unidad proposicional de esta índole tiene así hasta siete elementos componentes y podría derivar de una o más oraciones. Análogamente, GERBNER (1964) extrajo proposiciones a partir de programas informativos con el objeto de analizar orientaciones ideológicas que no se pondrían en evidencia en las palabras aisladas ni en la referencia a un objeto particular.

Unidades temáticas

Las unidades temáticas se identifican por su correspondencia con una definición estructural particular del contenido de los relatos, explicaciones o interpretaciones. Se distinguen entre sí sobre bases conceptuales, y del resto del material irrelevante por poseer las propiedades estructurales deseadas.

Las unidades temáticas exigen una profunda comprensión de la lengua fuente, con todos sus matices de significado y contenido. Aunque a menudo los lectores corrientes pueden reconocer fácilmente los temas, no les es tan fácil, por lo general, identificarlos de manera fiable. Aunque para juzgar qué clase de unidades son las más significativas es importante tener en cuenta el propósito de la investigación, puede que para muchos análisis de contenido las unidades temáticas sean las preferibles; no obstante, a raíz de la larga serie de operaciones cognitivas que conlleva su identifica-

ción, incluso los observadores mejor adiestrados pueden perderse con facilidad. Por lo tanto, habitualmente, el análisis de contenido evita las unidades temáticas, o, a lo sumo, las utiliza para circunscribir el confuso universo del cual se extraen la muestra o las unidades proposicionales.

Un buen ejemplo del empleo de unidades temáticas nos lo suministran KATZ y otros (1967), quienes analizaron cartas enviadas a los servicios administrativos del gobierno de Israel con el objeto de esclarecer en qué medida se hacía una buena utilización de estos servicios. Sus unidades temáticas se definieron como la solicitud de beneficios o exenciones a las autoridades, que incluían como elementos constitutivos descripciones de las calificaciones personales del autor de la carta y de las razones por las cuales pedía que se le concediese la prebenda.

Las unidades temáticas son de uso corriente en el análisis del folclore. La determinación de las unidades de los contenidos folclóricos se remonta a THOMPSON (1932), cuya lista y descripción de motivos ocupaba ya seis grandes volúmenes, y tenía como propósito la formulación de un esquema exhaustivo de registro. ARMSTRONG (1959) pasó revista a algunos de los problemas que plantea el uso de unidades temáticas en los contenidos folclóricos. Considerando la historia como una especie de folclore, el Council on Inter-racial Books for Children (1977) publicó una lista de temas en los que estaban presentes la discriminación sexual y racial, con el fin de identificar los estereotipos, deformaciones y omisiones recurrentes en los textos de historia norteamericanos.

EFICIENCIA Y FIABILIDAD

Las cinco unidades que acabamos de examinar se diferencian fundamentalmente por la clase de operaciones cognitivas que exige su identificación. En general, la determinación de unidades es tanto más eficiente y fiable cuanto más simples y "naturales" sean dichas operaciones cognitivas. Pero puede que las unidades simples no sean analíticamente productivas. El investigador, pues, tendrá que perfeccionar la productividad, sin perder demasiada eficiencia o fiabilidad.

Las unidades físicas exigen esencialmente un dispositivo mecánico. En este caso las operaciones cognitivas son mínimas, y por ello dichas unidades son eficientes y fiables. Sin embargo, a

menos que los límites de las unidades físicas coincidan con los del contenido que se quiere describir, pueden generar registros poco fiables, y dar lugar a hallazgos que carezcan de interés. Aunque establecer las unidades de muestreo a partir de las características físicas del medio tiene sus virtudes, rara vez las unidades de registro son definidas en estos términos.

Las unidades sintácticas exigen familiaridad con la gramática del lenguaje fuente, el medio de transmisión o la forma del material cuyas unidades se quieren determinar. La identificación de estas unidades suele ser eficiente y fiable, pero no siempre es productiva para el análisis posterior. Las unidades de contexto a menudo se definen de este modo.

Las unidades referenciales exigen conocer bien la semántica del lenguaje fuente, los símbolos y el significado referencial de los elementos. La identificación de estas unidades es bastante eficiente, pero no siempre fiable, y la dificultad principal radica en que las referencias no siempre son claras, salvo que las unidades se limiten a palabras o a breves frases denotativas. Estas unidades son las preferidas para la definición de las unidades de muestreo y de registro.

Las unidades proposicionales requieren un grado considerable de familiaridad con la sintaxis, la semántica y la lógica del lenguaje fuente, así como con ciertas transformaciones lingüísticas, como la reformulación de las frases, su realización, la descontextualización y la determinación de los núcleos de significado. Dado que la identificación de estas unidades exige con frecuencia reescribir todo un documento ateniéndose a un formato establecido, el proceso puede volverse bastante ineficaz y quizá sólo presente un grado moderado de fiabilidad.

En lo referente a la determinación de las unidades, puede establecerse como recomendación general la siguiente: debe apuntarse a las unidades que sean empíricamente más significativas y productivas, que puedan identificarse de manera eficiente y fiable, y que satisfagan los requisitos de las técnicas disponibles. Esto suele exigir transigencia en numerosos aspectos. A veces simplemente implica desprenderse de la información poco fiable estableciendo unidades proposicionales en lugar de temáticas, o referenciales en lugar de proposicionales.

6. Muestreo

Como en la mayoría de las actividades, el analista social debe recurrir a alguna clase de programa de muestreo con el fin de hacer posible su tarea. En este capítulo se exponen diversas estrategias de muestreo, particularmente aplicables a las necesidades del análisis de contenido.

La comunicación ha penetrado siempre en todas las esferas de la vida, pero en nuestros días la producción masiva de material impreso, la existencia de equipos para el registro del sonido y la imagen, las máquinas fotocopiadoras y los ordenadores electrónicos han aumentado enormemente la disponibilidad de material simbólico. Cuando un analista del contenido se plantea un problema para el cual puede haber respuesta en periódicos, películas cinematográficas, registros oficiales, etc., se encuentra ante un alud de información producida por diversas instituciones. Y cuando un investigador decide "simplemente grabar en videocassette" los experimentos grupales que lleva a cabo, debe tener en cuenta que el análisis de contenido de esas cintas fácilmente puede llevarle entre diez y cien veces el tiempo que abarca la proyección de la cinta en sí. El universo de datos primarios disponibles tiende a sobrepasar incluso la capacidad de asimilación de los operativos de investigación mejor equipados.

Enfrentado a esta cantidad de datos, el analista de contenido debe tomar dos clases de decisiones. En primer término, deberá emplear todos los conocimientos que pueda obtener con el fin de diferenciar el material relevante del que no lo es. La información conducente a obtener las inferencias pretendidas puede estar distribuida de forma desigual en distintos medios, publicaciones, documentos, períodos cronológicos o zonas geográficas. En segundo término, si una vez agotado este conocimiento disponible, el volumen de material relevante sigue siendo demasiado grande, tendrá que recurrir a métodos aleatorios para seleccionar una muestra que sea lo bastante amplia como para contener información suficiente, y lo bastante pequeña como para facilitar el análisis.

Supongamos que las inferencias relativas a la diferencia entre el actual presidente de un país y el candidato a sucederlo difieren en relación con su respectiva sensibilidad frente a las creencias y actitudes de la población. Ante todo, el investigador debe identificar las fuentes de información relevante. Los discursos grabados de ambos candidatos pueden servirle de punto de partida, pero tendrán que ser complementados con datos relativos a las características del público al que uno u otro candidato pudo haber tenido acceso, como las encuestas de opinión pública o las noticias periodísticas acerca de las expectativas efectivas del grupo de electores al que se dirigen, la manera en que votaron en pasadas elecciones o sus intereses e inquietudes. Pueden existir buenos motivos para efectuar el muestreo entre periódicos, circulares y declaraciones, de acuerdo con las preferencias de estos grupos de población en materia de medios de comunicación, ya que ellos, al igual que los candidatos, pueden llegar a conocerse entre sí a través de los medios que frecuentan. Estas decisiones comienzan a delinear el *universo de datos primarios* relevantes para las inferencias que se buscan.

Una vez establecido este universo de material relevante, deberán adoptarse decisiones acerca del segundo tipo de actividad, el *muestreo*. La necesidad práctica del muestreo procede de la reducción de una gran cantidad de datos potenciales a un tamaño manipulable. Su justificación metodológica deriva de que el proceso debe dar lugar a una muestra a partir de la cual puedan efectuarse generalizaciones seguras. Sin embargo, en el análisis de contenido se tiende a examinar una población o universo para formular generalizaciones acerca de otra; aquí el muestreo con-

cierne a dos poblaciones que, potencialmente, pueden ser muy diferentes.

Tenemos entonces la población* de observaciones posibles o universo de los datos que podrían obtenerse en caso de contar con recursos ilimitados. En nuestro ejemplo, esta población estaría constituida por todos los escritos sobre las expectativas políticas, referencias electorales, inquietudes públicas manifiestas y discursos posteriores a esas manifestaciones pronunciados en las campañas políticas de los candidatos. Pero, por otra parte, tenemos también el contexto al cual se dirigen las inferencias, que tiene que ver con la forma en que los oradores adecuarán sus expresiones, promesas y reflexiones de tipo valorativo a las características del público que ellos conocen. Este es el universo de las inferencias posibles. Extraer una muestra representativa de los datos posibles no es lo mismo que extraer una muestra representativa de lo que se ha de inferir. Un analista de contenido tiene que decidir qué clase de muestreo va a realizar.

TIPOS DE PLANES DE MUESTREO

Todos los procesos de muestreo están orientados por un *plan de muestreo*, que especifica con suficiente detalle de qué manera ha de proceder el investigador para obtener una muestra de unidades que, en su conjunto, sean representativas de la población que le interesa. Con el fin de obtener una muestra representativa, el plan asegura que, dentro de las limitaciones impuestas por el conocimiento disponible acerca de los fenómenos, cada unidad tiene la misma probabilidad de estar representada en el conjunto de unidades de muestreo. Esto garantiza que no haya tendenciosidad alguna en la inclusión de unidades en la muestra.

Muestra aleatoria

Suponiendo que no exista ningún conocimiento *a priori* sobre los fenómenos en cuestión, un plan de muestreo que tienda a obte-

* El término "población" procede del muestreo sobre poblaciones reales que han de ser entrevistadas. Aplicado a conjuntos de mensajes puede extrañar un tanto, pero en la perspectiva del autor equivale al conjunto de mensajes a estudiar. [E.]

ner una *muestra aleatoria* simple implicará el listado de todas las unidades relevantes (ejemplares de periódicos, documentos, discursos u oraciones) respecto de las cuales se pretende formular generalizaciones. Con el fin de determinar qué unidades habrán de ser luego incluidas en la muestra, el plan tal vez exija el uso de dados, una ruleta, una tabla de números aleatorios o cualquier otro dispositivo que adjudique iguales probabilidades a cada unidad. En el análisis de contenido se utiliza este procedimiento o cualquiera de los cuatro que a continuación mencionaremos.

Muestras estratificadas

El muestreo estratificado reconoce la existencia dentro de una población de varias subpoblaciones diferenciadas, a las que se denomina "estratos". Cada unidad de muestreo pertenece únicamente a un estrato. El muestreo aleatorio se lleva a cabo en cada estrato por separado, de modo que la muestra resultante refleja las distinciones que, *a priori*, se sabe que existen dentro de la población.

Por ejemplo, los periódicos han sido estratificados según su zona geográfica de distribución, la frecuencia de su publicación, su cantidad de lectores o la composición de su público (por ejemplo, lectores de periódicos prestigiosos en oposición a lectores de "prensa amarilla"). Otro ejemplo es la preparación de una semana típica de programas televisivos estratificando la programación de todo un año en días de la semana y horarios, y luego seleccionando al azar para cada horario una de las cincuenta y dos posibilidades existentes.

Muestreo sistemático

El muestreo sistemático implica la selección de cada unidad "k-ima" de una lista en la muestra, después de haber determinado al azar el punto de partida de la secuencia.

En el análisis de contenido, el muestreo sistemático se ve favorecido cuando los datos provienen de publicaciones de aparición regular, secuencias de interacción interpersonal, o del orden sucesivo de los escritos, películas cinematográficas y piezas musicales. El principal problema de esta clase de muestreo es que el intervalo de amplitud "k" es constante, y por ello crearía muestras tendenciosas si coincidiese con ritmos naturales como las varia-

ciones estacionales y otras regularidades cíclicas. Por esta razón, es conveniente no seleccionar los ejemplares de un diario cada siete números sino, digamos, cada cinco. La tendenciosidad que puede introducirse está ejemplificada en un estudio sobre anuncios matrimoniales en la edición dominical del *Times*, de Nueva York. Dicho estudio llegó a la conclusión de que entre 1932 y 1942 no habían existido anuncios de matrimonios en sinagogas judías (HATCH y HATCH, 1947); más tarde se comprobó que el muestreo sistemático de todos los ejemplares del periódico correspondiente a los meses de junio de esa década coincidió con un período durante el cual la tradición prohibía los casamientos entre los judíos (CAHNMAN, 1948).

Muestreo por conglomerados

El muestreo por conglomerados utiliza como unidades muestrales grupos de elementos que presentan designaciones y límites naturales. La selección de uno de estos grupos incorpora a la muestra todos sus elementos y, dado que los grupos contienen un número desconocido de ellos, la probabilidad de que una unidad sea incluida en la muestra dependerá del tamaño del grupo.

Si un investigador quiere estudiar la forma en que se presenta a los grupos minoritarios por la televisión, le será imposible enumerar o conocer por anticipado la población de los personajes televisivos, pero puede empezar por la lista de programas. De hecho, casi todo el material procedente de los medios de comunicación de masas viene en "paquetes" o "porciones" naturales (véase lo dicho sobre las unidades sintácticas): publicaciones regulares, noticias diarias, páginas o partes destinadas a la publicidad, libros, informes, historiales clínicos y cuentos de hadas son fácilmente identificables, por más que su contenido y composición varíen considerablemente. Pero si se seleccionan los programas televisivos dejándose llevar por la conveniencia, se incorporarán a la muestra una cantidad desigual de personajes. Esto ocurre también cuando se seleccionan los periódicos por ejemplares dentro de los cuales se identifican, computan o comparan un número variable de noticias divididas en apartados. En realidad, siempre que las unidades de muestreo y las unidades de registro no sean equivalentes, puede estar implícita, subrepticiamente, la formación de un conglomerado.

Este tipo de muestreo constituye una respuesta práctica frente a

la imposibilidad de enumerar de manera individual los elementos de una población, cuando sí se pueden enumerar los grupos en los que aparecen. Pero los resultados del agrupamiento en conglomerados están casi siempre sujetos a mayores errores de muestreo. La variancia dentro de la muestra tiende a alcanzar valores excesivamente altos. El uso de conglomerados es preferible a la selección individual de los elementos sólo cuando el ahorro de esfuerzo por cada elemento es mayor que el aumento de variancia.

Muestreo de probabilidad variable

En el muestreo de probabilidad variable se asignan las probabilidades de inclusión en la muestra de cada unidad de acuerdo con algún criterio *a priori*. El procedimiento más común es el submuestreo en el que las probabilidades se asignan de acuerdo con el tamaño, lo que da lugar a las que suelen denominarse *muestras proporcionales*. Los criterios de asignación de probabilidades deben ser explícitos (a diferencia de lo que ocurre en las muestras por conglomerados) y justificarse en función del diseño global de la investigación.

El muestreo de probabilidad variable es importante para el análisis de contenido debido a que se dedica a la formulación de inferencias sobre fenómenos no incluidos en la muestra (ni observados directamente, ni muestreados). Como ejemplo hipotético, supóngase que un analista de contenido quiere inferir opiniones de la población rusa acerca de los Estados Unidos y se ve limitado a utilizar solamente periódicos. Al planear un estudio así, debe tenerse en cuenta que los comentarios que aparecen en los periódicos son muestras sistemáticamente "tendenciosas" de las opiniones de la población. En primer lugar, no todos los grupos de la población tienen el mismo acceso a los medios de prensa, y en segundo lugar, algunas opiniones tienen más probabilidad de ser divulgadas que otras. Una muestra que quisiera ser representativa de todos los periódicos existentes incorporaría esta "tendenciosidad" directamente en los datos, y haría extraer inferencias carentes de validez acerca de las opiniones de los individuos. Si se pretende confeccionar una muestra de periódicos a partir de la cual puedan hacerse inferencias sobre la población de individuos, deben anularse los efectos de este proceso de automuestreo. Esto puede efectuarse asignando a cada opinión manifiesta una probabilidad que guarde una relación inversa con la probabilidad de que esa opinión llegue

a divulgarse por la prensa. Así, las declaraciones de portavoces oficiales aparecerán con menos frecuencia en la muestra que las manifestaciones publicadas de habitantes comunes.

MACCOBY y otros (1950) expusieron una interesante aplicación de este procedimiento al análisis de contenido de periódicos. Estos autores se interesaron en los periódicos no como institución, sino por la información a que tenían acceso los lectores. Hicieron una lista de todos los diarios publicados en nueve distritos censales (estratos), añadieron las cifras correspondientes a sus tiradas respectivas y asignaron a cada periódico una probabilidad de acuerdo con su participación en la tirada total. Las unidades de muestreo eran los ejemplares de periódicos, mientras que la cantidad de lectores determinaba las probabilidades que se les asignaban.

Las probabilidades variables representan el conocimiento estadístico que posee un investigador acerca del contexto de los datos, o sea, acerca de la forma en que los fenómenos que le interesan están indicados probabilísticamente en los datos que en efecto se obtienen. Dado que ese conocimiento es a menudo incierto o hipotético, es difícil justificar el uso de muestras de probabilidad variable, y sólo deben extraerse con extrema precaución.

Muestreo en etapas múltiples

A menudo se obtienen muestras recurriendo sucesivamente a uno o más procedimientos de muestreo. A este procedimiento se le denomina *muestreo en etapas múltiples*, y puede considerarse una variante del muestreo por conglomerados. Por ejemplo, puede obtenerse una muestra de periódicos después de haberlos estratificado por zonas geográficas y cantidad de lectores, luego identificar sistemáticamente determinados temas dentro de cada periódico, y seleccionar más adelante artículos correspondientes a esos temas asignándoles probabilidades proporcionales a su extensión. El muestreo en etapas múltiples no siempre utiliza diferentes procedimientos en cada etapa; para obtener una idea sobre la bibliografía acerca de un determinado tema, puede partirse de una serie de artículos recientes sobre ese tema, extraer una muestra aleatoria de la bibliografía citada en los artículos y luego una segunda muestra aleatoria de las referencias encontradas en esa bibliografía citada, hasta que la inclusión de nuevos elementos no modifique la composición proporcional de la muestra.

Aunque parezca existir una considerable libertad para elegir cualquier plan de muestreo, el objetivo fundamental es obtener muestras representativas de los fenómenos que interesan. Por lo general, es más difícil justificar esta elección en el análisis de contenido que en la investigación por encuestas, de la cual proceden la mayoría de las técnicas y la terminología del muestreo, puesto que los fenómenos que interesan sólo se manifiestan de forma indirecta en el material disponible. En el análisis de contenido no son los datos el objeto de análisis, sino que éstos más bien constituyen un escalón de apoyo para avanzar hacia otros fenómenos.

Tamaño de la muestra

Una vez decidido el modo de extraer la muestra, el próximo problema es, habitualmente, determinar su tamaño. No existe para esto una solución preestablecida. Si todas las unidades muestrales son exactamente idénticas, una muestra constituida por una sola unidad será satisfactoria; esto es lo que a menudo se presume en ingeniería y en las encuestas destinadas a verificar los gustos de los consumidores. Si, por el contrario, en la lista de unidades hay algunos casos extraordinarios, infrecuentes y significativos, la muestra tendrá que ser amplia, y deberá incluir a la población total si cada unidad de muestreo es idiosincrásica. En la práctica, esta incertidumbre no es tan abrumadora: aunque cada unidad adicional de la muestra aumenta la dificultad del análisis, hay un punto en que ningún aumento posterior mejorará apreciablemente la generalizabilidad de los hallazgos. Ese es el punto en el que resulta más eficiente el tamaño muestral. Es una cuestión relacionada con la dificultad-beneficio, que depende en gran medida de la forma en que se distribuye en la muestra el atributo que se desea generalizar.

Stempel (1952) comparó muestras constituidas por seis, doce, dieciocho, veinticuatro y cuarenta y ocho números de un periódico con los números aparecidos a lo largo de todo un año, y comprobó, utilizando como medida la proporción media de materias tratadas, que el aumento del tamaño de la muestra más allá de doce no producía resultados significativamente más precisos. Es de lamentar que estas recomendaciones se limiten a estudios que utilizaron medidas similares y las aplicaron a periódicos que, a su vez, tenían una distribución similar del contenido.

Una prueba relacionada con el tamaño apropiado de una muestra que no exige analizar la población total es la llamada *técnica de la división en mitades*, en la cual se divide aleatoriamente a una muestra en dos partes de igual tamaño; si cada una de las partes permite extraer las mismas conclusiones estadísticas, dentro del mismo nivel de confianza, puede aceptarse que la muestra total tiene un tamaño adecuado. Esta prueba puede repetirse en distintas divisiones equivalentes, ya que debe conservar su vigencia para tantas divisiones como requiera el límite de confianza. Si la prueba falla, el investigador deberá aumentar el tamaño de la muestra hasta que se satisfagan las condiciones.

7. Registro

Al crear datos a partir de observaciones o de un texto, es importante tener en cuenta las características de los codificadores, su capacitación, la sintaxis y semántica del lenguaje de datos utilizado y la forma en que se administra el procesamiento de los datos. En este capítulo se explican los procedimientos para desarrollar instrucciones adecuadas, y se suministran los correspondientes ejemplos.

El registro constituye uno de los problemas metodológicos fundamentales de las ciencias sociales y de las humanidades. La proposición, aceptada en las ciencias naturales, de que la realidad no es accesible como tal sino a través de la mediación de un instrumento de medición, debe aplicarse también en este caso. No es posible analizar lo que no ha sido adecuadamente registrado, ni puede esperarse que un material que sirve de fuente venga expresado ya en los términos formales de un lenguaje de datos. El registro es indispensable siempre que los fenómenos que interesan o bien carecen de estructura con respecto a los métodos disponibles, o bien son simbólicos, en el sentido de que son portadores de información sobre otros fenómenos que están más allá de sus manifestaciones físicas. En las formas simbólicas espontáneas abundan las omisiones y estructuras confusas, las ambigüedades, y las dependencias del contexto, que los instrumentos físicos de medición rara vez pueden registrar. Las comunicaciones simbóli-

cas reflejadas en escritos, registros sonoros y videocassettes deben ser en gran medida transcritas en términos formales antes de poder utilizarse para el procesamiento de datos y la inferencia.

En el análisis de contenido, el registro ha sido un problema tan importante que las antiguas definiciones equiparaban prácticamente aquél con el proceso de registro. Por ejemplo:

> Puede definirse el "análisis de contenido" diciendo que esta expresión alude a cualquier técnica a) para la clasificación de *vehículos-signos,* b) que se basa exclusivamente en los *juicios* de un analista o grupo de analistas en cuanto a las categorías a las cuales pertenece cada vehículo-signo (juicios que teóricamente pueden abarcar desde diferenciaciones perceptuales hasta meras conjeturas), c) siempre y cuando dichos juicios del analista se consideren como el informe de un observador científico (JANIS, 1965, pág. 55).

En otra caracterización leemos:

> Con el fin de manejar grandes conjuntos de material verbal de manera estadística, parece indispensable reducir la variedad de alternativas que deben disponerse en tablas; y esto puede lograrse situando en una sola categoría una amplia variedad de estructuras de palabras diferentes (MILLER, 1951, pág. 95).

El registro es una consecuencia necesaria del hecho de que el análisis de contenido acepte material no estructurado, pero no debe confundirse con el análisis de contenido del cual forma parte.

Ya nos referimos antes, en términos más generales, a que las instrucciones explícitas deben contener *todo* lo necesario para reproducir el proceso de elaboración de los datos, cuando intervienen diferentes individuos. La exigencia de Janis acerca de que las instrucciones sean explícitas se refiere primordialmente a la semántica del lenguaje de datos, o sea, a las reglas para asignar vehículos-signos a las categorías. Las *instrucciones de registro* explícitas deben comprender:

- las características de los observadores (codificadores, jueces) que intervienen en el proceso de registro;
- la capacitación y preparación que estos observadores reciben con el fin de prepararse para la tarea;
- la sintaxis y la semántica del lenguaje de datos utilizado, incluyen-

do en caso necesario un esbozo de los procedimientos cognitivos que se han de emplear para dividir los mensajes en categorías;
- la administración de las planillas de datos.

OBSERVADORES

Desde luego, los observadores, codificadores y jueces deben estar familiarizados con la naturaleza del material que han de registrar, pero además deben ser capaces de manejar fiablemente las categorías y términos que componen el lenguaje de datos. No es fácil cumplir con este doble requisito. Si el problema consiste en registrar un dialecto local, por ejemplo, es fácil encontrar hablantes nativos, pero mucho más difícil es encontrar alguno de ellos con la suficiente capacitación en el método científico. Por lo tanto, los analistas de contenido suelen abordar el problema desde el otro extremo, utilizando individuos con una cierta formación en ciencias sociales (por ejemplo, estudiantes universitarios) y que estén familiarizados con los fenómenos que se han de registrar. Aunque probablemente en estos casos las deformaciones sean menos notables, lo cierto es que un estudiante universitario común de clase media inevitablemente tendrá dificultades para apreciar las sutilezas de un dialecto o jerga de algún grupo minoritario local. Aun cuando se trate de describir espectáculos de televisión, temas y personajes a los que tienen acceso muchas personas, los antecedentes socioeconómicos y lingüísticos, así como las diferencias educativas entre los codificadores, suelen ser decisivos en cuanto a la posibilidad de registrar fiablemente los datos.

Así pues, las instrucciones sobre registro deben incluir una caracterización de la clase de individuos a los que están destinadas dichas instrucciones y con respecto a los cuales se ha verificado la fiabilidad de éstas. No puede presumirse simplemente que las instrucciones de registro carecen de ambigüedad para todo el mundo y que, por lo tanto, todos pueden cumplimentarlas exactamente de la misma manera. En general, los problemas prácticos de evaluación de la reproducibilidad son menos graves si las instrucciones se dirigen a observadores

- que abundan dentro de la población
- respecto de los cuales es fácil afirmar que cumplen con el doble requisito antes mencionado.

CAPACITACIÓN

La capacitación de los codificadores es una labor preparatoria corriente en el análisis de contenido. No sólo los individuos tienen que estar familiarizados con las peculiaridades de la labor de registro (rara vez los procedimientos y definiciones se ajustan perfectamente a la intuición), sino que además estos codificadores a menudo desempeñan un papel instrumental en la plasmación del proceso, sobre todo en la fase preparatoria del análisis de contenido. Lo normal es que los investigadores informen que las sesiones de capacitación del personal se prolongaron durante varios meses, y que en ese período las categorías se refinaron, los procesos se modificaron y las planillas de datos se revisaron hasta que los individuos se sintieron cómodos y estuvieron en condiciones de efectuar su tarea de forma fiable y eficiente. A veces las ideas primitivas llegan a modificarse hasta tal punto que ya no pueden reconocerse. Un buen ejemplo del modo en que las definiciones del lenguaje de datos surgen de forma paralela a la capacitación de los codificadores se describe en un estudio sobre las actitudes de la política exterior soviética y norteamericana.

> La finalidad del estudio era trazar un cuadro preciso sobre los objetivos y estrategias de soviéticos y norteamericanos en materia de política exterior, en tanto y en cuanto éstos pudieran reflejarse en las manifestaciones de la elite sobre: A) el ambiente internacional; B) la distribución del poder, C) el código operativo del otro país, y D) el código operativo propio.
>
> El procedimiento se dividió en dos fases principales: la de diseño y refinamiento de nuestro procedimiento de codificación, y la de su aplicación. A su vez, en la primera fase se siguieron seis etapas más o menos delimitadas:
>
> 1. Se compilaron las preguntas que parecían más afines al estudio que se tenía entre manos. Por supuesto, éstas se basaron en una multiplicidad de fuentes: el conocimiento general que el autor tenía sobre el tema, los parámetros de los esquemas conceptuales de su propia ciencia social, y las dimensiones de la política exterior sugeridas por los escritos e investigaciones de otras personas que trabajaron en este mismo campo.
> 2. Una vez establecida y ordenada una serie provisional de dimensiones esencialmente *a priori*, éstas fueron examinadas, criticadas y modificadas por el autor, sus ayudantes, algunas personas a quienes se acudió para consultar y otros colegas profesionales.

3. Los codificadores aplicaron luego esta serie de dimensiones a una muestra del material que debía codificarse, lo cual dio como resultado la supresión de algunas dimensiones, la reformulación de otras y el añadido de unas pocas nuevas.
4. A continuación el autor reevaluó las dimensiones y redujo aún más las tres categorías que abarcaba cada una de ellas, con el objeto de llevar al máximo su exclusión recíproca así como la exhaustividad de dichas categorías.
5. Luego las dimensiones y sus categorías se sometieron a un pretest a cargo de los propios codificadores, para asegurar que:
   a. En la bibliografía que habría de codificarse se hacía referencia a las dimensiones que valía la pena codificar.
   b. Las dimensiones mismas no se superponían entre sí (salvo en unos pocos casos, en que se procuraba captar algunos sutiles matices de actitud).
   c. Las dimensiones en sí mismas eran lo bastante claras e inequívocas como para garantizar la existencia, entre codificadores independientes, de un alto grado de acuerdo en cuanto a qué artículo específico debía codificarse o no en esa dimensión.
   d. Las tres categorías alternativas en el interior de cada dimensión eran tan excluyentes entre sí como había sido posible establecer, pese a lo cual eran exhaustivas con respecto a las posibles gamas de respuestas relevantes.
6. Cuando los pretests demostraron (por acuerdo entre dos o más codificadores independientes) que las dimensiones y categorías ya habían sido refinadas y aclaradas adecuadamente, se consideraron definitivas. (SINGER, 1964, págs. 11-12.)

Estas etapas son típicas, por lo siguiente:

- El diseñador de la investigación formula sus requisitos iniciales en lo que se refiere a los datos.
- Se familiariza con la forma en que la información relevante se expresa en el material que ha de tomar como fuente.
- Formula instrucciones de registro por escrito.
- Trabajando en común con los codificadores que han de aplicarlas, las instrucciones son interpretadas y modificadas entre todos hasta que satisfacen los requisitos adecuados de fiabilidad.

La calibración tiene lugar en la última etapa, y en el ejemplo anterior el problema radica en que está implícita. En este proceso se alcanza algún grado de homeostasis entre lo que pretende la persona que diseñó la investigación, lo que ven los observadores y

lo que pueden aprender, lo que implican las instrucciones y la forma en que puede interpretarse el material que sirve de fuente sin violentar la intuición. Como consecuencia de ello, sólo quienes han participado en este proceso de ajuste mutuo podrán operar de forma congruente. Las interpretaciones específicas de cada grupo modifican las instrucciones escritas, que entonces dejan de ser exclusivamente representativas del proceso de registro. Resumiendo el uso del análisis de contenido en psicoterapia, LORR y MCNAIR (1966, pág. 583) apuntan los efectos que estos procesos implícitos de ajuste tienen sobre la reproducibilidad:

> Aunque la mayoría de los investigadores publican admisibles índices de acuerdo entre los codificadores en la categorización de las respuestas, dichos índices están sujetos a serios cuestionamientos. Por lo general, el índice de acuerdo entre codificadores que sale a la luz se basa en dos personas que han trabajado juntas íntimamente en el desarrollo del plan de codificación, y que han debatido durante mucho tiempo las definiciones y sus discrepancias al respecto. Un acuerdo entre calificadores correspondiente a un nuevo grupo de jueces, con un período razonable aunque práctico de capacitación con un sistema, representaría un índice de fiabilidad más realista. Las pruebas que se llevaron a cabo con algunos sistemas vigentes de análisis de contenido sugirieron que la fiabilidad obtenida mediante un nuevo grupo de jueces, utilizando sólo las reglas de codificación formales, definiciones y ejemplos, era muy inferior con respecto a lo que se informa habitualmente. Con frecuencia no satisfacían las normas mínimas del trabajo científico.

Lo ideal es que los individuos que toman parte en el desarrollo de instrucciones de registro adecuadas no participen en el registro mismo de los datos. Una vez establecidas dichas instrucciones, cualquiera que sea el método utilizado, debería añadirse una quinta etapa a la lista anterior:

- La fiabilidad de las instrucciones de registro se verifica aplicándolas a un nuevo grupo de observadores independientes.

Los individuos deben poder operar con las instrucciones de registro como única guía. Sólo deben tener un acceso mínimo, si es que tienen alguno, a las fuentes de información no controlables (por ejemplo, la historia del proceso por el cual se llegó a esas instrucciones), ya que de lo contrario las instrucciones podrían modi-

ficarse en direcciones imprevisibles; además, deben poder operar con un grado absolutamente mínimo de comunicación informal entre ellos, para que no surjan acuerdos subrepticios sobre sus propias interpretaciones. Cualquier modificación de las instrucciones que sea necesario introducir deberá incorporarse en las propias instrucciones escritas.

Si se exige capacitación, es preciso que sea estandarizada, para poder reproducirla en otros sitios. En cierta oportunidad establecimos un minucioso programa de autoaprendizaje para el registro de episodios de violencia televisiva: se impartió sintéticamente a los sujetos la naturaleza de la tarea, y a continuación se les dejó trabajar solos, observando una serie corriente de programas de televisión. Una vez que identificaban y registraban cada unidad en una planilla de datos, los sujetos podían encontrar las puntuaciones supuestamente correctas (establecidas por un equipo de expertos) en otra planilla. La comparación entre las dos planillas les proporcionaba una realimentación inmediata sobre su propio trabajo y les facilitaba adoptar una interpretación estándar de las instrucciones. Este método no sólo nos permitió trazar un esquema de fiabilidad creciente, sino también decidir, al término del período de aprendizaje, cuáles eran los individuos más adecuados para esta tarea. Si se cuenta con un programa de autoaprendizaje de esta índole, el proceso es fácil de reproducir, y casi forzosamente produce siempre resultados similares.

Probablemente una de las peores prácticas difundidas en materia de análisis de contenido sea el hecho de que el investigador desarrolle sus instrucciones de registro y las aplique solo, o con ayuda de unos pocos colegas íntimos, impidiendo así someterla a controles independientes de su fiabilidad. Suele aducirse que esta costumbre obedece a la falta de recursos, pero una prueba tras otra han demostrado que este proceso carece en gran medida de fiabilidad, y en realidad nos lleva a preguntarnos qué clase de contribución científica puede realizar un estudio que sólo el autor es capaz de reproducir.

SEMANTICA DE LOS DATOS (MANERAS DE DEFINIR EL SIGNIFICADO DE LAS CATEGORIAS)

La *sintaxis* y la *semántica* de un lenguaje de datos están incorporadas esencialmente a las reglas que gobiernan la asignación de

unidades a las categorías o al código. Las marcas efectuadas en una planilla de datos, los orificios de una tarjeta perforada, las crípticas anotaciones que hace un analista al margen del texto fuente, son portadores de información en la medida en que se sigan las reglas de manera fiable; pero sólo resultan significativos desde el momento en que las propiedades que dan origen a los datos estén representadas reconociblemente en ellos. Los datos son entidades simbólicas.

Las instrucciones de registro no sólo deben asegurar que los datos se registren de una manera fiable, sino también explicar su significado. Sólo cuando la relación semántica entre los datos puntuales y el material que sirve de fuente es clara, los hallazgos basados en esos datos pueden conducir a intelecciones acerca de los fenómenos reales. Si se pierden las instrucciones de registro o los manuales de codificación, por más que los datos exhiban su sintaxis, las inferencias serán inciertas porque el investigador ya no tendrá acceso a su semántica. A menudo, la semántica de los datos se contamina simplemente por una codificación poco fiable, o por modificaciones introducidas en las instrucciones mientras se llevaba a cabo la codificación, o por el uso de definiciones operacionales especiales, que luego se equiparan con el concepto intuitivo que tal vez les dio origen. El ruido, las ambigüedades y las perspectivas contrarias a la intuición en la semántica de un lenguaje de datos hacen más difícil la subsiguiente interpretación de los hallazgos.

Con frecuencia se enuncia el requisito de que "las categorías sean exhaustivas y mutuamente excluyentes". Este requisito corresponde a la semántica del lenguaje de datos, por cuanto establece una relación entre los fenómenos que se han de describir y los datos que los representan. La "exhaustividad" se refiere a la capacidad de un lenguaje de datos para representar todas las unidades de registro, sin excepción. Ninguna unidad debe quedar excluida porque se carezca de términos descriptivos adecuados. La propiedad de "exclusión mutua" se refiere a la capacidad del lenguaje de datos para establecer distinciones netas entre los fenómenos que se han de registrar. Ninguna unidad debe pertenecer a dos categorías o estar representada por dos datos puntuales distintos. Este doble requisito exige que la semántica de un lenguaje de datos divida el universo de unidades de registro posibles en *clases bien diferenciadas* y que los miembros de cada una de estas clases estén representados por un *dato diferente*, de modo que las

distinciones establecidas en el universo aparezcan inequívocamente representadas en los datos.

Un conjunto de categorías que no es exhaustivo puede llegar a serlo si se le añade otra categoría que represente a todas las unidades no representadas en el conjunto inicial. A menudo esas categorías se etiquetan como "otros", "el resto", "no aplicable", etc. Dado que una categoría "autoprotectora" de esta índole no representa un conjunto claramente delimitado de fenómenos, salvo por exclusión de todos los restantes, contribuye poco (si es que en algo contribuye) a los hallazgos de la investigación y, por ello, debería evitarse en lo posible recurrir a ella.

Esta solución no es aplicable a categorías que carezcan de la propiedad de exclusión mutua. Añadir una categoría "ambigua" al conjunto, una categoría en la que se reúnan todas las unidades con respecto a las cuales resulta poco clara la asignación de categorías, o que puedan adoptar valores múltiples (como cuando hay dos o más categorías aplicables), impide evaluar la fiabilidad, y lo que es más importante, orienta tendenciosamente los resultados de la investigación en la dirección de los fenómenos fácilmente descriptibles. La ambigüedad de las instrucciones de registro no tiene remedio.

El modo en que se definen las categorías y en que se adoptan valores numéricos o datos puntuales representativos de los fenómenos reales, de las observaciones y de las características de los mensajes constituye todo un arte, sobre el cual se ha escrito muy poco. Sin embargo, pueden diferenciarse algunas *maneras de delinear la semántica* de un lenguaje de datos:

- designaciones verbales
- listas de extensión
- esquemas de decisión
- magnitudes y escalas
- simulación de la verificación de hipótesis
- simulación de entrevistas
- construcciones conceptuales de cierre e inferencia

Designaciones verbales

Las designaciones verbales son las más frecuentes y, en apariencia, las más obvias. Por ejemplo, el sexo de un personaje dramático puede ser o bien *masculino*, o bien *femenino*, o bien *inde-*

terminado. Esta última designación puede obedecer a que no se cuenta con la información suficiente, a que se trata de un niño sin un rol sexual claro, o a algún otro motivo. Pero las designaciones de características o atributos con una sola palabra, incluyendo los nombres de individuos, conceptos o clases de fenómenos, sólo son eficaces cuando las diferenciaciones entre ellas se ajustan plenamente a los significados lingüísticos corrientes. Y esto puede no suceder, ya que toda disciplina científica desarrolla invariablemente sus propias concepciones teóricas. El siguiente conjunto de categorías ilustra distinciones que no se comparten con observadores no avezados o, por citar un ejemplo, con los pacientes psiquiátricos cuyo nivel de angustia se procura evaluar (MAHL, 1959), en cuyo caso se requieren evidentemente definiciones y ejemplos más amplios.

1) "¡Ah!" Cada vez que se presenta con claridad la interjección "¡ah!", se la registra.
2) *Corrección de oración* (CO). Toda corrección en la forma o contenido de la expresión a medida que avanza la secuencia de las palabras emitidas. Para que se registren, el oyente debe percibir estos cambios como una interrupción de dicha secuencia.
3) *Oración incompleta* (OI). Casos en que una expresión cualquiera se interrumpe, se deja incompleta, prosiguiendo la comunicación sin que el hablante haga ninguna corrección.
4) *Repetición* (R). La repetición superflua en la serie de una o más palabras (habitualmente sólo una o dos).
5) *Tartamudeo* (T).
6) *Interferencia de un sonido incoherente* (SI). Un sonido que es absolutamente incoherente como palabra para el oyente, y que interfiere sin alterar la forma de la expresión, y sin que pueda interpretarse claramente como un tartamudeo, una omisión o un desliz verbal (aunque de hecho algunas de estas incoherencias lo sean en la realidad).
7) *Desliz verbal* (DV). En esta categoría se incluyen los neologismos, la trasposición de palabras respecto de su posición correcta dentro de la serie y la sustitución de la palabra que se quería por otra involuntaria.
8) *Omisión* (O). Pueden omitirse partes de palabras o, lo que es más raro, palabras completas. Se exceptúan de esta clasificación las contracciones. En su mayoría, las omisiones son sílabas terminales de palabras.

Listas de extensión

Una lista de extensión especifica la semántica de un lenguaje de datos indicando para cada término del material que se toma como fuente la categoría a la que pertenece. Aunque en la construcción de esta lista puede subyacer un esquema conceptual, no es necesario que el codificador lo conozca. Las listas de extensión constituyen un requisito en determinados enfoques del análisis de contenido mediante el uso de ordenadores. Por ejemplo, el programa General Inquirer incluye un procedimiento llamado el "diccionario", mediante el cual se establece la equivalencia de una gran cantidad de términos del texto de entrada con respecto a un pequeño conjunto de palabras o expresiones típicas.

Las listas de extensión son necesarias siempre que faltan convenciones lingüísticas y diferenciaciones conceptuales, y son ventajosas cuando no resulta sencillo comunicar estas últimas a los codificadores. Véase el ejemplo de O'SULLIVAN (1961), que intentó cuantificar la intensidad de las presuntas relaciones entre las variables en los escritos teóricos sobre las relaciones internacionales. La adopción previa del análisis factorial requería que la "intensidad de la relación" se conceptualizara como una correlación; pero al capacitar a los codificadores, pronto se puso en evidencia que las designaciones verbales de la idea subyacente no conducían a diferenciaciones fiables; de ahí que se impusiera el uso de una lista:

1. Es menos probable que; en ciertas situaciones provoca; puede originar algunos; puede obedecer a; puede ser.... en la medida en que; puede aplicarse sin; las posibles consecuencias que se desprenden de ello parecen ser.
2. Ha introducido un elemento adicional; no es simplemente función de... sino de; es un factor de; no sólo depende de... sino de; depende en parte de; es posible que provoque.
3. Origina; es probable que sea; suele producir; tendería a; tenderá a producir; tiende a; tiende a introducir.
4. Convierte en improbable que; afecta fuertemente a; es muy probable que derive de; es muy probable que acontezca; crea esencialmente; depende primordialmente de; es uno de los principales motivos de; crea un problema del tipo de.
5. Realzará; exige como mínimo; subrayará; necesita; determinará; produce; depende de; es inevitable; es el resultado de; reflejará; impondrá; impide; sobrepasará; debilita; fortalece; ofrece el máximo de; será menor que; se añadirá a.

6. Cualquiera de ellos debe antes; son mínimos cuando; si ...entonces; tiene; establece; es; es menos ...cuando ha...; si es ...entonces es; existe; ha existido, y existe; ha sido, y es; está directamente relacionado con; se intensificará en relación directa con; guarda una relación inversa con; influirá en... en proporción directa a; está en relación directa con; hay una relación directa entre; se halla en marcado contraste con; en la medida en que; cuanto más...tanto más; cuanto mayor... tanto mayor; cuanto mayor... tanto más; cuanto mayor... menor; cuanto mayor... y mayor... tanto mayor; cuanto mayor... más; cuanto más amplio... menor; cuanto más... menos; cuanto más... más; cuanto más..., y mayor... más; cuanto más... mayor; cuanto más... menos probable es que; más... que; cuanto más amplio... mayor; cuanto más amplio... más; cuanto más alto... mayor; cuanto más largo... menor; cuanto más corto... mayor debe ser; cuanto menos sean... mayor será; se vuelve más... a medida que; tiene más probabilidades de ser... cuanto más; son menos... cuanto menos; son menos... si son menos; serán más... cuanto más grande; cuanto más grande... más;

Esquemas de decisión

Los esquemas de decisión consideran cada dato como el resultado de una secuencia predefinida de decisiones. Su empleo tiene varias ventajas importantes. En primer término, con ellos pueden evitarse los problemas que surgen cuando las categorías poseen diferente nivel de generalidad o cuando su significado se superpone. SCHUTZ (1958) ejemplificó esto con un estudio de historietas, clasificando de la siguiente manera los lugares en que se desarrollaba la acción: el territorio de los Estados Unidos, el territorio de países extranjeros, zonas rurales o urbanas, lugares históricos y el espacio extraterrestre. En la figura 8 se reordenan estas categorías como consecuencia de varias decisiones dicotómicas, siendo las categorías terminales las del lenguaje de datos, y las restantes, simplemente conceptos intermediarios.

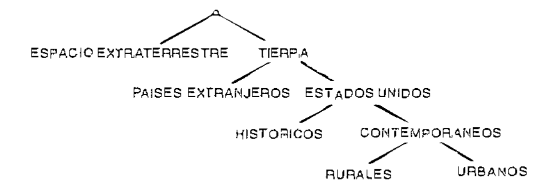


Figura 8. Esquema de decisiones para el registro.

En segundo término, cuando las unidades de registro son pluridimensionales, los esquemas de decisión ofrecen la oportunidad de descomponer una enunciación compleja en varias decisiones simples, alcanzando de ese modo niveles de fiabilidad que no podrían obtenerse de otra manera.

En tercer lugar, los esquemas de decisión reducen de forma drástica la cantidad de alternativas que deben examinarse de forma simultánea en cada etapa.

Magnitudes y escalas

Las magnitudes y escalas exigen que el codificador conceptualice el material que le sirve de fuente como un continuo, o como si estuviera ordenado o poseyera una métrica propia. Las escalas de diferencial semántico propuestas por OSGOOD y otros (1957) pueden servir de ejemplo:

fuerte	:——:——:——:——:——:——:——:	débil
activo	:——:——:——:——:——:——:——:	pasivo
bueno	:——:——:——:——:——:——:——:	malo

Cada escala queda conceptualmente establecida por dos términos opuestos polares y, sin necesidad de explicar los significados de los puntos intermedios de la escala, se le pide al codificador que describa los atributos de cada unidad de registro puntualizando a qué distancia se encuentra de esos dos extremos polares. Dicho sea de paso, estas tres escalas corresponden a lo que, según se ha comprobado, constituyen las tres dimensiones básicas de la cognición afectiva humana: la potencia, la actividad y la dimensión evaluativa (SUCI, 1957; OSGOOD 1962).

No todos los conceptos pueden situarse en un continuo entre opuestos. Para la ley, los atributos "legal-ilegal" pueden ser opuestos binarios, mientras que la distinción "noticias-ficciones" no es evidentemente unidimensional. Si se pretende imponer una escala a algo que no se presta a esta concepción, puede darse origen a medidas poco fiables. En realidad, en muchos análisis de contenido las escalas de diferencial semántico resultan poco fiables principalmente a causa de que se carece de información sobre el atributo que debe registrarse, o a que dicha información es insuficiente. Por ejemplo, en algunas ficciones televisivas ciertos personajes son descritos pormenorizadamente, mientras que otros

apenas aparecen en una escena breve. Como es lógico, las escalas bipolares de la personalidad son más fiables para los personajes principales que para los secundarios. A todas luces, los juicios acerca del grado en que se suministra información sobre una característica se encuentra en un nivel diferente de los juicios acerca de lo que esa característica es en sí misma. ZILLMAN (1964) introdujo una escala que elude esta dificultad, con el nombre de "escala del aspecto semántico". Consta de seis puntos, en los cuales la presencia de una determinada característica se interpreta así: "ausente-muy poco-poco-mediano-mucho-muchísimo". Esta escala unipolar resulta apropiada cuando los atributos, cualidades o fenómenos pueden estar presentes en mayor o menor medida, o ser más o menos importantes, más o menos intensos, o más o menos frecuentes, y exigen concebir al continuo como un coeficiente (con un cero absoluto).

Simulación de la verificación de hipótesis

En su mayoría, las expresiones verbales llevan consigo presupuestos e implicaciones lógicas que trascienden los atributos y referencias que establecen. Por ejemplo, la frase "El leía el *New York Times*" presupone que el sujeto en cuestión posee una visión suficiente, que está en un lugar adecuadamente iluminado, que conoce la lengua inglesa y tiene por lo menos seis años de edad; también implica que ha recibido alguna información acerca de lo que es "publicable" para ese periódico. Aunque probablemente sea imposible preparar una lista completa de estas inferencias, no es tan difícil relacionarlas cognitivamente con enunciados o hipótesis predefinidos. En general, una hipótesis es un enunciado cuya verdad se rechaza por un ejemplo contrario, por refutación o por pruebas estadísticas que favorecen una hipótesis opuesta. Como estrategia de registro, la simulación de la verificación de hipótesis exige al codificador relacionar de manera cognitiva y lógica cada unidad de registro verbal con cada una de las distintas hipótesis mutuamente excluyentes, y cerciorarse de a cuál corresponde y, a veces, en qué medida apoya o rechaza una u otra alternativa.

Un ejemplo clásico de este tipo de simulación se encuentra en el método adoptado por LASSWELL (1965b) para detectar la propaganda extranjera en las fuentes de noticias nacionales durante la segunda guerra mundial. Lasswell extrajo cuatro finalidades básicas de la propaganda a partir de los pronunciamientos públicos del

partido nazi y de los funcionarios del gobierno alemán, y solicitó a sus codificadores que juzgaran si los nuevos ítems y comentarios acerca de los acontecimientos apoyaban o rechazaban implícitamente alguna de estas finalidades, o varias de ellas.

La simulación de la verificación de hipótesis mediante procesos cognitivos-lógicos tienen usos que trascienden la mera indicación de la tendencia. Gracias a ella es posible evaluar el grado en que la descripción de fenómenos psicológicos se ajusta a una u otra teoría sobre el comportamiento humano, o la manera en que se explican los accidentes, o las concepciones populares que están en la base de los debates sobre las enfermedades y los tratamientos médicos, o el tipo de teorías de la comunicación que un autor da por sentado implícitamente cuando informa acerca de una investigación.

Simulación de entrevistas

La simulación de entrevistas proporciona un medio de obtener datos semejantes a los de una encuesta mediante los escritos de un "entrevistado". Se solicita al codificador que se familiarice con la obra del autor, formándose una imagen mental de su personalidad, conocimientos y creencias, para más tarde, repasándola por segunda vez, buscar datos que indiquen si el autor podría responder sí o no, o dar alguna otra respuesta matizada intermedia, a un conjunto de preguntas previamente preparadas que un entrevistador le haría si pudiera.

Un ejemplo de una entrevista simulada de este tipo lo da KLAUSNER (1968), que basó sus inferencias en una muestra estratificada de 199 manuales de puericultura, de un total de 666 publicados en los Estados Unidos en un período de dos siglos. Las actitudes y concepciones de cada autor se registraron en función de ochenta preguntas con respuestas predefinidas. En la página siguiente se incluye un ejemplo de una de esas preguntas.

Si la simulación de la verificación de hipótesis descansa fundamentalmente en la capacidad lógica y lingüística del codificador, la de las entrevistas descansa en una habilidad adicional: la de asumir el papel del autor. Naturalmente, esto añade fuentes de pérdida de fiabilidad, en particular si el volumen de escritos de los autores es amplio y poco específico. Cuanto más breves sean los escritos, más específicas deben ser las preguntas, con el fin de evitar caer con excesiva asiduidad en la categoría "sin respuesta" u

Términos del lenguaje de datos		Items del cuestionario y respuestas en lenguaje fuente
Nº de columna	*Perforación*	
32		*Pregunta:* ¿Cómo legitima el libro la autoridad del progenitor ante este último? (¿En qué se basa su apelación al progenitor para que preste atención al niño?)
		Respuestas:
	1	No lo examina.
	2	Da por sentada la legitimación, pero no indica ninguna base específica.
	3	El progenitor tiene mayores conocimientos que el niño.
	4	El progenitor es moralmente superior al niño (apelación al sentido de responsabilidad personal).
	5	El progenitor es un representante moral de la comunidad.
	6	El progenitor influye moral e intelectualmente en el niño, lo quiera o no, y por ello es responsable de las consecuencias de sus propios actos.
	7	El progenitor influye psicológicamente en el niño, lo desee o no.
	8	Otros.
	0	NA (no aplicable, no se ocupa de esta cuestión).

"otros". Las ventajas de las simulaciones de esta índole son que los entrevistados no reaccionan frente al entrevistador, que resulta admisible plantear preguntas embarazosas (BARTON, 1968) y que los entrevistados están "disponibles", siquiera de forma indirecta, por un período indefinido.

Construcciones conceptuales de cierre e inferencia

Una de las marcas características de los especialistas en psicodiagnósticos es que pueden examinar lo que un paciente evita decir, además de lo que expresa abiertamente. De manera similar, los analistas políticos tienden a "leer entre líneas" y los críticos culturales procuran identificar pautas en lo que está sospechosamente ausente. Estos trabajos pueden asemejarse al "cierre de una figura incompleta". Abandonados a una intuición ilimitada, estos cierres se volverían, desde luego, sumamente subjetivos y proyectivos; pero es posible, al menos en ciertas ocasiones, especificar de antemano "la organización abstracta de la figura" o las construcciones cognitivas que han de emplearse para obtener cierres fiables.

Encontramos un ejemplo en las elucidaciones de GEORGE (1959a) sobre las inferencias extraídas por la FCC norteamericana a partir de las emisiones radiofónicas internas del enemigo durante la segunda guerra mundial. En el curso de su labor, los analistas desarrollaron elaboradas construcciones conceptuales

con el propósito de explicar por qué aparecieron ciertas emisiones radiofónicas y cuáles fueron las percepciones y condiciones previas. En este ejemplo no resultaba tan clara la distinción entre el registro de los datos y la formulación de inferencias. Mencionaremos con más detalle este enfoque en el capítulo sobre construcciones analíticas; aquí basta con decir que los analistas crearon y emplearon construcciones sumamente específicas acerca de la situación, incluidas algunas generalizaciones sobre el comportamiento político y la conducta en materia de propaganda de la elite gobernante enemiga, lo cual les permitió obtener una considerable cantidad de información militar.

Los procesos cognitivos a que recurren estos codificadores para llegar a las conclusiones deseadas los describe muy bien GEORGE (1959, pág. 61):

> La argumentación del analista consiste en el establecimiento de cada una de las principales variables inestables que aún se desconocen (o de asignarles un valor), apoyando esta reconstrucción en generalizaciones, y en evaluaciones de la lógica de la situación. Este tipo de razonamiento inferencial puede equipararse al trabajo de reconstruir las piezas ausentes de un mosaico. Algunos fragmentos del mosaico aparecen ya determinados, o resulta fácil presuponerlos, en tanto que faltan otros, incluidas las condiciones que al analista le interesa particularmente esclarecer. En consecuencia, lo que en la práctica hace el analista es ensayar mentalmente las diferentes versiones posibles de cada variable ausente que pretende inferir, procurando decidir cuál es la versión más admisible, dado el valor conocido de la variable de contenido y los valores conocidos o postulados de otras condiciones antecedentes.

Esta enumeración de métodos para operacionalizar la semántica de un lenguaje de datos no es en modo alguno completa: aquí sólo hemos enunciado los tipos principales, de modo que puedan servir de base para nuevos y mejores métodos de registro.

Aunque existen diferencias entre estos métodos con respecto a la fiabilidad y la eficiencia, es imposible pronunciarse en favor de cualquiera de ellos. Sabemos que las listas de extensión breves son mejores que las largas, pero en cambio el punto en el cual la codificación mediante designaciones verbales se vuelve superior a la lista de extensión depende de tantas circunstancias que ninguna generalización simple puede ayudar a quien diseña la investigación. Como regla general, el registro es más fiable y eficiente

cuanto más familiares resulten los conceptos, más simples sean las operaciones cognitivas, más breves las cadenas inferenciales, menos cuantiosas las categorías, y cuanta menos capacitación requieran los codificadores.

Planillas de datos

Las planillas de datos contienen información en su forma más primaria y explícita. Si los fenómenos que se desean registrar están en curso y no dejan huellas materiales, como los debates grupales y los programas de televisión en directo, estas planillas pueden constituir los *únicos* registros de esos fenómenos; y si además los fenómenos de interés no están estructurados, como sucede con los discursos grabados en cintas magnetofónicas o como suele ocurrir con las cartas manuscritas, las planillas de datos contienen *toda* la información, acerca de esos fenómenos, que puede tener en cuenta un análisis de contenido. Quizá la expresión "planilla de datos" no sea lo bastante general, ya que algunos datos primarios pueden registrarse directamente sobre cinta magnética (Janda, 1969) o filmados (Ekman y Friesen, 1968), aunque nuestro material más común de registro es el papel.

El diseño de registro primario puede exigir bastante ingenio. Como las exigencias que se plantean son tan variadas, es imposible sugerir una forma estándar u óptima; no obstante, pueden hacerse unas pocas recomendaciones. El requisito de que todas las unidades de registro de la misma índole se describan en función de la misma serie de categorías, sumado a la necesidad de contar con técnicas de duplicación y de impresión, favorece sin lugar a dudas la separación de las planillas de datos por unidades de registro y no según los observadores o según las variables. De este modo, las unidades de registro del mismo tipo deben apuntarse en la misma clase de planillas de datos, que puedan luego duplicarse o imprimirse. Un análisis de contenido exige tantos registros separados como unidades de registro.

Las planillas de datos contienen varias especies de información, algunas de ellas obvias, otras fácilmente soslayables. En síntesis, pueden reducirse a las tres especies siguientes:

- información administrativa;
- información sobre la organización de los datos;

- información sobre los fenómenos que se desea registrar, o sea, los datos en sí mismos.

Información administrativa

La información administrativa orienta el manejo burocrático de los datos. La necesidad de espacio para esta clase de información no debe subestimarse. Con frecuencia ocurre que las planillas de datos desordenan la secuencia y hacen imposible su reconstrucción. Se pierde mucho tiempo si, frente a una serie de formularios completados, no se puede saber con certeza a qué parte del estudio corresponden, o si ya han sido verificados, perforados o almacenados.

La información administrativa incluye:

a) La identificación del *proyecto de análisis de contenido* para el cual se registran los datos, incluida la fase de desarrollo del instrumento de registro, si éste ha sido modificado, y tal vez la identificación del tipo general de los datos y de la clase de unidad a la que se aplican. Esta información puede imprimirse en cada planilla, y sirve para diferenciar entre los distintos proyectos o fenómenos que se están analizando.

b) La identificación del *estado de la planilla de datos*: completada, verificada, archivada (por número), duplicada, transferida a otra planilla, perforada, etc. Sea cual fuere el modo de procesamiento de la planilla, el director de la investigación debe ser capaz de discernir cómo se ha manejado y qué hay que hacer aún con ella antes de archivarla.

c) La identificación de *los individuos participantes* en el manejo de los datos, particularmente el observador que codificó los fenómenos originales, pero también aquellos que verificaron y procesaron los datos, y a los que quizás haya que consultar si más adelante se encuentran errores.

d) Instrucciones respecto de *la manera en que se han de perforar los datos* (número de columna para las variables, número de las categorías) o cualquier otro procedimiento que se utilice para preparar las planillas con el fin de que los ordenadores las procesen.

En algunos estudios, la cantidad de información administrativa tal vez parezca desproporcionadamente grande, pero es indispensable.

Información sobre la organización de los datos

La información sobre la organización de los datos será importante siempre que se empleen en el análisis varias clases de unidades de registro. Para utilizar una clase de datos en el análisis de otros, debe añadirse dicha información, ya sea en forma de archivo principal que indique de qué manera se ensamblan entre sí las diferentes clases de datos, o en forma de entradas especiales que indiquen cuál es la correspondencia mutua entre los datos.

Por ejemplo, en un estudio sobre artículos editoriales periodísticos, puede destinarse una clase de planilla de datos al periódico en que apareció el artículo (e incluir allí su tirada o su frecuencia de publicación), otra clase de planilla para el contenido del artículo y una tercera para los ítems informativos asociados con éste. Repetir simplemente la información sobre los periódicos en que se publicaron dichos editoriales para cada uno de éstos que se analice no sólo exige tiempo, sino que además impide un análisis separado de los propios periódicos (en lugar de los artículos editoriales) como unidad de enumeración. En este caso, números de referencia arbitrarios, que correspondan al periódico en que se publicó el editorial y al ítem informativo al que se alude en éste, permiten acceder a distintos tipos de datos. Análogamente, en un análisis de contenido de las pautas de interacción entre personajes televisivos, registraremos en una clase de planilla los rasgos de la personalidad y otras características estables e idiosincrásicas del personaje, mientras que incorporaremos los contactos comunicativos a una matriz de interacción, con números de referencia correspondientes a los personajes. Esto nos permitirá comparar las características diferenciadoras de los individuos.

Datos

Por supuesto, la razón de ser del proceso de registro en su conjunto es la *información sobre el fenómeno que se desea registrar*, es decir, *los datos propiamente dichos*. El primero y más obvio requisito es que los observadores puedan anotar fácilmente esta información, y que puedan también leerla fácilmente quienes

deben procesarla, pero que en cambio no resulte sencillo alterarla, ya sea por desgaste o destrucción material, o con algún propósito deshonesto. Los aparatos de exploración óptica exigen utilizar un tipo especial de lápices, lo cual hace más difícil verificar las marcas. Cuando los datos van a ser perforados en tarjetas, las planillas a menudo se organizan de manera que las entradas del manual puedan traducirse sin complicaciones en el formato de ochenta columnas de la tarjeta perforada. Si el análisis se lleva a cabo manualmente, el uso de algunas abreviaturas contribuirá a una mejor comprensión.

Aparte de esta facilidad para el ingreso y lectura de los datos, la forma de las planillas debe ser tal que reduzca al mínimo los errores. Una fuente típica de error es el uso de entradas ilegítimas. Examínense las tres maneras siguientes de registrar el resultado de las interacciones interpersonales:

Coloque el número apropiado	Rodee con un círculo sólo una opción		Controle cuál de las dos opciones se aplica
☐ 0 - favorable a ninguno de los dos	*favorable a*	ninguno de los dos	☐ favorable al iniciador
1 - favorable al destinatario		destinatario solamente	
2 - favorable al iniciador		iniciador solamente	☐ favorable al destinatario
3 - favorable a ambos		ambos	

Las diferencias entre estas tres formas de registro radican fundamentalmente en el tipo de errores a que pueden dar lugar. En el caso de la versión de la izquierda, se pueden anotar números mayores que tres, los cuales no están definidos en ningún lugar y por lo tanto resultan ilegítimos; además, esta forma depende más que las otras de las peculiaridades de la escritura a mano. En la versión del medio se insta a utilizar dobles códigos, en tanto que en la de la derecha no se toleran en absoluto marcas ilegítimas. Aunque no siempre es posible prevenir del todo la aparición de entradas ilegítimas, las planillas de datos pueden prepararse de modo que estos errores sean mínimos.

Otra fuente principal de errores está relacionada con la necesidad de repetir en las planillas de datos la semántica del lenguaje de datos. Como las instrucciones de registro pueden olvidarse con facilidad, un caso extremo sería repetir todas las definiciones de las categorías; esto asegura altos niveles de consistencia, pero exige mucho tiempo y es costoso. En el otro extremo, sólo se presenta al codificador una gran cuadrícula en la cual debe insertar las entradas numéricas; esto provoca a veces múltiples confusiones de filas y columnas, y desconcierto en cuanto a lo que sig-

nifican en cada caso estas entradas numéricas. Hemos comprobado la utilidad del siguiente procedimiento: 1) incluir en las planillas una indicación de lo que significa cada variable, y preferiblemente cada opción; 2) utilizar signos de puntuación, y evitar los números y letras cuando éstos no tengan una relación intrínseca con los fenómenos que se desea registrar; y 3) emplear la particular organización de la planilla de datos para transmitir con ella algunas de las características organizativas de las unidades de registro. Esto puede aplicarse concretamente al registro de las propiedades icónicas, la distribución espacial, la secuencia temporal ("antes" y "después" pueden convertirse en "izquierda" y "derecha", respectivamente), así como a las distinciones conceptuales entre el emisor, el mensaje y el receptor.

8. Lenguajes de datos

Las categorías y mediciones, que son las intermediarias entre el mundo de los fenómenos reales y el de los hechos científicos, condicionan el éxito del análisis. En este capítulo se examina un sistema de categorías y mediciones como lenguaje de datos. Su semántica relaciona un dato con las observaciones y mensajes, mientras que su sintaxis lo relaciona con el procedimiento científico.

El aparato descriptivo en cuyos términos un analista traduce sus datos se denomina *lenguaje de datos.* Todo lenguaje de datos es un intermediario entre el mundo de los fenómenos reales y el de los hechos científicos, y con frecuencia se convierte en un obstáculo para las intelecciones científicas. Un lenguaje de datos posee una sintaxis y una semántica. *La semántica de un lenguaje de datos vincula un dato cualquiera con el mundo real, mientras que su sintaxis lo vincula con el método científico.*

El problema de definir los significados operacionales de las categorías de un análisis constituye el eje principal del *registro.* La recomendación ambivalente de extraer las categorías ya sea del material que se toma como fuente, ya sea de teorías relevantes, se plantea en la distinción émica-ética. Y la falta de ambigüedad en la designación de los fenómenos reales se considera al ocuparse de la *fiabilidad.* En este capítulo examinaremos la *sintaxis de los lenguajes de datos* que son importantes para el análisis de contenido. En general, los lenguajes de datos:

- deben estar libres de toda ambigüedad e incongruencias sintácticas;
- deben satisfacer las demandas formales que plantea la técnica analítica para ser aplicable;
- deben poseer una capacidad descriptiva que les permita suministrar, acerca de los fenómenos que interesan, información suficiente como para resultar concluyentes.

El primero de estos tres requisitos exige que los lenguajes de datos estén formalizados. En principio, sólo los lenguajes formales o formalizados son computables. Los seres humanos corrientes, al ser sensibles al contexto, están bien capacitados para hacer frente a las ambigüedades sintácticas, pero las técnicas analíticas explícitas no lo están. Por ejemplo, si un lector común lee la oración "están volando allá arriba", y tiene acceso al contexto, no tendrá dificultad alguna para adivinar si "están" se refiere a un grupo de pilotos aéreos o bien a diversos objetos que se ven en el cielo. Para el procesamiento mediante ordenador, será necesario eliminar esta ambigüedad sintáctica, ya sea mediante procedimientos especiales o mediante algún añadido editorial que identifique qué acepción del verbo "volar" se utiliza aquí. La oración "Vienen Jim o Joe y Mary" es análogamente ambigua, y exige algún añadido editorial que distinga "(p o q) y r", de "p o (q y r)". [En este caso, para determinar si los que vienen son "(Jim o Joe) y Mary", o bien "Jim o (Joe y Mary)".]

Como regla general, a menos que un análisis haga caso omiso de las características sintácticas (como sucede en los cómputos de palabras, en los que la posición de éstas carece de importancia), las expresiones en lenguaje natural tienen que codificarse y transcribirse, o al menos editarse, con el fin de que se conviertan en sintácticamente inequívocas y puedan ser computables. *La sintaxis de un lenguaje de datos debe estar bien definida.*

El segundo requisito tiene en cuenta el hecho de que cada técnica analítica plantea sus propias exigencias con respecto a la forma de los datos. Esto puede parecer obvio, pero es sorprendente la frecuencia con que los investigadores comprueban que no pueden satisfacer los requisitos del análisis una vez reunidos los datos. Por ejemplo, los datos sobre hábitos de lectura y la descripción del contenido de un periódico, analizados de forma separada, no permiten explicar las preferencias de un lector en

materia de contenido. El motivo es que las técnicas de correlación y de asociación exigen que los datos sean registrados en pares. En este caso, cada lector tendría que formar una pareja con el material al que tiene acceso. Los análisis en series cronológicas requieren que el objeto de análisis sea observado en diferentes momentos. Los datos sobre una muestra de objetos en un momento y los datos sobre una muestra de objetos en otro momento no satisfacen este requerimiento. Otro error común surge cuando las variables de un estudio no poseen una métrica lo suficientemente adecuada como para aplicar la técnica. Los análisis de variancia requieren que los datos estén expresados en intervalos; un ordenamiento por rangos descalificaría a los datos para un análisis de este tipo. Cuando los datos no satisfacen los requisitos de una técnica analítica determinada, en general los resultados no pueden interpretarse (una concepción opuesta al respecto sostiene TUKEY, 1980).

El tercer requisito procede del objetivo del análisis de contenido, es decir, de las inferencias que se pretende lograr. LASSWELL (1960) caracterizó en cierta oportunidad las investigaciones sobre la comunicación mediante la siguiente pregunta: "¿Quién dice qué, por qué canal, a quién y con qué efecto?", estableciendo que el análisis de contenido se especializa en la pregunta "¿Quién dice qué?", mientras que las investigaciones sobre el público conciernen a "¿A quién?". Una vez realizadas estas distinciones, sin embargo, no supo ver que la respuesta separada a las preguntas "¿Quién?", "¿Qué?", "¿A quién?", "¿Con qué efecto?", no permite obtener intelecciones acerca de la comunicación como proceso social mediativo. Para estudiar este proceso se necesita un lenguaje de datos lo bastante eficaz como para rastrear el flujo de la información como un fenómeno pluridimensional (KRIPPENDORFF, 1970). Los lenguajes de datos pueden fallar en lo referente a ofrecer la información suficiente, ya sea adoptando una perspectiva desde la cual no sea posible abarcar la totalidad del conjunto (como en el caso de la separación de Lasswell entre el análisis de contenido y el análisis de otras facetas del fenómeno de la comunicación), dejando fuera variables importantes, o estableciendo un número insuficiente de distinciones. En todos estos casos, el efecto es que se omiten detalles estructurales significativos. Con frecuencia, los requisitos en materia de información —ya sea para poner a prueba la validez de las hipótesis científicas, para evaluar el éxito de las acciones programadas o para disipar la incertidum-

bre acerca de un problema— pueden especificarse de antemano. *Un lenguaje de datos debe proporcionar como mínimo toda la información que exige el objetivo de un análisis de contenido.*

Teniendo en cuenta estos tres requisitos, podemos definir un lenguaje de datos en términos lo bastante generales como para dar cuenta de todas las inquietudes de los analistas de contenido.

DEFINICION

Un lenguaje de datos prescribe la forma en que deben registrarse los datos, y consiste en:

- *variables* cuyos valores representan la variabilidad de las unidades de registro dentro de una dimensión conceptual;
- *constantes* con significados operacionales fijos que especifican las relaciones entre las variables;
- una *sintaxis* cuyas reglas gobiernan la construcción de registros bien conformados (fórmulas, expresiones) a partir de las variables y las constantes;
- una *lógica* que determina qué registros se implican recíprocamente o deben considerarse equivalentes; establece dependencias lógicas (*a priori*) entre las variables.

En la fórmula algebraica

$$ax + b = c,$$

"a", "b", "c" y "x" son variables, cada una de las cuales puede asumir un valor numérico; el signo "+" y el signo de multiplicación implícito en la expresión "ax" son constantes que designan operaciones algebraicas bien definidas; el signo "=" es un signo lógico que designa que las dos partes de la fórmula son equivalentes y mutuamente reemplazables. Además, los dos miembros de esta fórmula están bien conformados. En cambio, de acuerdo con las reglas del álgebra, la serie de símbolos "a b x c = +", por ejemplo, no estaría bien conformada. En el proceso de registro, las variables se sustituyen por sus valores particulares. Obviamente, no todas las combinaciones de valores satisfacen esta fórmula, de modo que la lógica y la sintaxis fijan una restricción en cuanto a la gama de configuraciones admisibles de valores.

Una sintaxis excluye ciertas combinaciones de los valores que pueden adoptar las variables, considerándolas ilegítimas, mientras que *una lógica equipara ciertas combinaciones* de estos valores (convierten en superfluas algunas distinciones). Ambas reducen la variación que una serie de variables puede registrar a las que efectivamente son necesarias.

En el análisis de contenido, la sintaxis y la lógica de los lenguajes de datos suelen ser muy simples o estar ausentes, en cuyo caso los datos se registran en su *forma básica* o más primaria. Cada unidad de registro se describe en función de un número fijo de variables, cuyos valores se insertan en casilleros como los que muestra la figura 9. En este caso, las constantes no hacen sino asignar una posición arbitraria a cada variable.

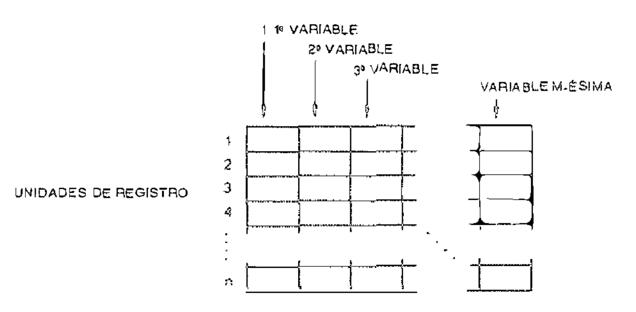


Figura 9. Forma básica de disponer los datos en m dimensiones en el análisis de contenido.

Mencionamos aquí la sintaxis y la lógica de un lenguaje de datos principalmente porque se están produciendo avances notables en otros campos, particularmente en la lingüística, que deben tenerse en cuenta en el análisis de contenido. Por ejemplo, las gramáticas transformacionales cuyas sintaxis incluyen reglas de reescritura que tienden a caracterizar las expresiones del lenguaje natural no pueden representarse mediante un formato con un número fijo de variables (o sólo pueden serlo de una manera muy grosera). Incluso análisis de contenido menos ambiciosos pueden incluir formas recursivas. Así. en un análisis de contenido sobre las actitudes de los nativos y los extranjeros en Africa (PIAULT, 1965), las respuestas a preguntas abiertas se registraron en los siguientes términos:

a) Una serie ordenada de variables. concernientes a características sociales del entrevistado.

b) El origen del juicio, por ejemplo:

X juzga que Y es ( )
X se refiere a Y que juzga que X es ( )
X se refiere a Y que se refiere a X…

c) Relaciones entre las X y las Y, con respecto al origen del juicio.
d) Temas, es decir, tres clases de argumentos asociados con cada juicio.
e) Un léxico, compuesto por las variables (675 términos) y las constantes (operadores lógicos "y", "o", y otros semejantes) que pueden utilizarse para presentar lexicalmente cada argumento.

En este caso, b) incluye una definición recursiva de un registro bien conformado, mientras que e) no puede representarse mediante un número fijo de variables. El análisis de los datos que llevó a cabo Piault se apoyaba en rutinas de recuperación de información cuyas demandas sintácticas satisfacían las de este lenguaje de datos. Aunque las reglas sintácticas y la lógica de un lenguaje de datos son importantes, esto es todo cuanto podemos decir por el momento. Las consideraciones siguientes se ocupan del uso de las variables y su ordenamiento.

VARIABLES

Una variable es un símbolo que representa un valor cualquiera dentro de una serie de dos o más valores mutuamente excluyentes, como objetos, estados, categorías, cualidades o elementos. WEBSTER (1967) define una variable diciendo que es "algo capaz de variar o está en condiciones de hacerlo, algo modificable". La variación es lo que permite que los datos sean *informativos*. En realidad, si no existieran hombres *y* mujeres, la variable sexo carecería de todo significado descriptivo; si los periodistas no tuvieran la opción de inclinarse en favor de uno *u* otro bando en una polémica, la medida de la "desviación" tampoco tendría sentido. En un análisis de contenido, cada unidad de registro debe poder caracterizarse mediante uno de varios valores alternativos respecto de cada variable, y la elección de dicho valor excluye todos los restantes. Esto nos lleva a otras maneras de definir las variables: *una variable divide en clases mutuamente excluyentes el conjunto de todas las unidades de registro.*

Como las nociones de "variable" y de "valor" son generales, pueden expresarse con distintos términos; he aquí algunas correspondencias:

variable -	valor
serie -	elemento
conjunto -	miembro
dimensión -	punto
escala -	puntuación
conjunto de la categoría -	categoría
espacio -	localización
calibración -	medición
sistema -	estado
tipología -	tipo

Cuando los analistas de contenido describen los problemas referidos a la construcción de categorías, o cuando los psicólogos construyen escalas de calificación de siete puntos, o cuando los especialistas en teoría de los sistemas definen un sistema por los estados que éste adopta, en todas esas circunstancias está subyacente la cuestión de "variable".

Las variables pueden ser *abiertas* o *cerradas*. La edad es una variable abierta pues sus valores no tienen un límite superior lógico.

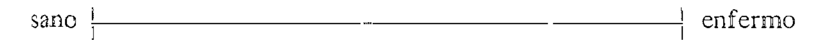
El estado civil, en cambio, es una variable cerrada, ya que las categorías tienen un número limitado y son conocidas de antemano. Aunque siempre se recomienda definir la variable con respecto a una única dimensión conceptual, las variables cerradas no necesitan ajustarse a las dimensiones convencionales, en la medida en que den lugar a valores o categorías bien diferenciados. En general, todo investigador tiene mayor libertad conceptual para definir las variables cerradas que las abiertas.

Una lista completa de valores proporciona una definición explícita de una variable. Una definición implícita es ejemplificada por un esquema de codificación que exige otro tipo de reformulación del texto:

( ) dice ( ) a ( ).

Por ejemplo: (Jim) dice ("hola") a (Mary). En este caso, los nombres y los mensajes reformulados se introducen como los

valores o categorías de las tres variables, implícitamente definidas por constantes relacionales. Son también abiertas, porque estas variables pueden corresponder a los nombres de dos comunicadores cualesquiera y a cualquier mensaje que ellos intercambien. Por el contrario, una escala de opuestos polares es también una variable definida implícitamente, pero cerrada:



Si en una escala de esta índole pueden identificarse muchos puntos, por numerosos que sean, tienden a escogerse mediante alguna función intuitiva de las distancias que los separan de uno y otro extremo. No resulta tan fácil saber qué métodos de ordenamiento o esquemas tienen en mente los codificadores.

Las variables pueden asimismo derivarse, o sea, definirse, sobre la base de otras variables (primarias), abstraer de estas últimas ciertas características o tomar la forma de índices que, por lo tanto, ignoran una parte de la variación indeseable de la variable original. Por ejemplo, pueden registrarse las veinte posibles "redes de comunicación", en un grupo de cinco personas, como una sola variable, en oposición al registro de la presencia o ausencia de los diez posibles "eslabones de comunicación" entre los pares de miembros de ese mismo grupo. O bien puede computarse la velocidad en lugar de la rapidez, o combinar varias medidas distintas en una puntuación única. El índice de violencia de GERBNER y otros (1979) y la medida de la atención propuesta por BUDD (1964), por ejemplo, permiten obtener esas medidas derivadas o secundarias definiéndolas a partir de registros primarios o de combinaciones de valores de ciertas variables.

Los analistas de contenido suelen sacar a la luz su "esquema conceptual" o su "sistema de categorías" sin dejar bien claro cuáles han sido, de hecho, las variables utilizadas en el estudio. Véase el siguiente esquema (HERMA y otros, 1943), tal como se describe en la bibliografía:

Criterios para rechazar la teoría de los sueños de Freud

A. Subestimación a través de juicios valorativos.
   1. Ridiculización y burla.
   2. Rechazos basados en cuestiones morales.
   3. Negación de la validez.

B. Impugnación del carácter científico de la teoría.
   1. Cuestionamiento de la sinceridad del analista.
   2. Cuestionamiento de la verificación de la teoría.
   3. Cuestionamiento de la metodología.
C. Exposición de la condición social de la teoría.
   1. Discrepancia entre los especialistas.
   2. "Está de moda".
   3. Falta de originalidad.

Una interpretación posible es que aquí sólo hay una variable, que puede adoptar nueve valores (A1, A2, hasta C3), siendo las divisiones A, B y C simples agrupamientos de estos valores en tres categorías menos precisas. Una segunda interpretación es que hay tres variables (A, B y C), de las cuales 1, 2 y 3 actúan respectivamente como valores; en este caso podría presumirse que cada argumento contrario a la teoría de los sueños de Freud posee una dimensión valorativa, una dimensión científica y una dimensión social. La tercera interpretación es que hay aquí nueve variables cuyos valores son "presente" o "ausente", y que la partición en A, B y C no es más que una guía conceptual para el codificador. Una clave para descifrar de qué manera se definen las variables radica en la variación que pueden representar; una segunda clave radica en la enumeración de los datos. Como cada variable debe dividir el conjunto de todas las unidades de registro, sin excepción, la suma de las frecuencias asociadas con los valores de cada una debe ser igual al tamaño de la muestra. En este caso, de hecho, la interpretación utilizada fue la primera. Las frecuencias relativas de los nueve valores suman 100% y el agrupamiento permite que los hallazgos sean representados también en función de tres valores.

Escalas nominales

Las escalas nominales consisten en una serie de dos o más valores que no poseen ni un orden ni una métrica. Las escalas nominales son las formas básicas de las variables, puesto que sólo requieren la diferenciación de sus valores. Los datos registrados en las categorías nominales se denominan también "cualitativos", porque la diferencia entre dos valores cualesquiera de una escala nominal es la misma para todos los posibles pares de valores.

Entre los ejemplos de escalas nominales podemos mencionar las listas de nombres de individuos, países, ciudades, ocupaciones,

grupos étnicos, clases de palabras, tipos de expresiones o argumentaciones. Su ordenamiento alfabético nada tiene que ver con las unidades enumeradas. Por más que se usen números para identificar los valores de una escala nominal (como al designar a los jugadores de un equipo de fútbol o a los billetes de lotería), las diferencias numéricas entre estos valores no tienen significado en la escala. Otra forma de enunciar esta propiedad es decir que en el interior de una escala nominal las distinciones se mantienen aun cuando sus valores experimenten toda clase de permutaciones.

Para establecer una distinción entre variables distinta de la de las escalas nominales, empleamos estos dos conceptos:

- orden
- métrica

El orden está relacionado con el ordenamiento de los valores, con la red de relaciones existentes entre ellos. La diferencia entre una organización jerárquica y una uniforme reside en su ordenamiento. La métrica tiene que ver con la clase de operaciones matemáticas aplicables a estas relaciones ordenadas. Dos valores que representen los ingresos de un individuo pueden sumarse, restarse o multiplicarse, mientras que los atributos que distinguen a un fabricante de calzado de un fabricante de velas no admiten estas operaciones. La diferencia es de métrica. Todas las distinciones formales entre las variables se reducen a responder a tres interrogantes relacionados con las diferencias entre las unidades de

Tabla 1
Tipos de variables según orden y métrica

orden: métrico	agrupamientos	cadenas	circuitos cerrados	cubos	árboles	reticulados de partición
ausente	escala nominal					
ordinal	clasificación	escala ordinal	ciclo	tabulación cruzada multivariada	tipología	
intervalos		escala de intervalos	π (p. ej. tiempo)	espacio		
razón		escala de razones		espacio vectorial absoluto		

registro que están representadas por los valores de la variable. Si estas diferencias entre las unidades de registro son significativas para un estudio, los valores de una variable deben reflejarla; y si las relaciones entre los valores de una variable poseen las propiedades adecuadas, las técnicas analíticas no deben distorsionarlas en el curso del análisis.

En este contexto sólo se establecen unas pocas distinciones formales; en la tabla 1 se mencionan las que revisten especial importancia para el análisis de contenido.

Orden

Dado que cada unidad de registro es descrita en función de un (y sólo de un) valor de cada variable, toda relación existente entre dos o más de estas unidades debe estar representada implícitamente. Ejemplos de esas representaciones implícitas son las redes cognitivas de los conceptos utilizados por un autor (Baldwin, 1942), la red semántica de un texto tal como es almacenado en un ordenador (Klir y Valach, 1965) y el esquema de asociaciones que puede utilizarse para responder preguntas, o la jerarquía de funcionarios en una organización. En beneficio de la simplicidad, tómese a título de ejemplo el mapa vial del Sistema Amtrak de itinerarios ferroviarios para pasajeros. Al especificar de qué manera se relacionan entre sí, directa o indirectamente, las ciudades, lugares de interés e intersecciones, incluidas las rutas alternativas, los circuitos o los puntos terminales, la red ordena el conjunto de todos los puntos mutuamente excluyentes en un mapa. Teniendo en cuenta sólo ciertas estructuras o una característica cada vez, el análisis de esas redes suele considerarlas más simples de lo que realmente son. En el análisis de contenido, los datos suelen registrarse en formas muy simples, y analíticamente muy manejables. Para información de los diseñadores de investigaciones, a continuación delinearemos seis tipos comunes de orden: agrupamientos, cadenas, circuitos cerrados, cubos, árboles y reticulados de partición.

Agrupamientos

Los agrupamientos dentro de las variables indican que los valores de un grupo tienen más cosas en común que los de grupos

diferentes. Las categorías de WHITE (1951) para el análisis de valores personales se disponen en ocho grupos, cada uno de los cuales contiene entre uno y seis valores. He aquí otra clasificación en dos niveles, de tipo psicológico:

A. *Factores fisiológicos*
   1. Alimentación
   2. Relaciones sexuales
   3. Descanso
   4. Salud
   5. Seguridad
   6. Comodidad

B. *Factores sociales*
   1. Amor erótico
   2. Amor familiar
   3. Amistad

C. *Factores yoicos*
   1. Independencia
   2. Realización
   3. Reconocimiento
   4. Autoconsideración
   5. Dominio
   6. Agresividad

D. *Factores relacionados con el temor*
   1. Seguridad emocional

E. *Factores lúdicos*
   1. Nuevas experiencias
   2. Entusiasmo
   3. Belleza
   4. Humor
   5. Autoexpresión creativa

F. *Factores prácticos*
   1. Espíritu práctico
   2. Posesiones materiales
   3. Trabajo

G. *Factores cognitivos*
   1. Conocimiento

H. *Varios*
   1. Felicidad
   2. Valores en general

Este agrupamiento sugiere que la diferencia entre "alimentación" y "relaciones sexuales" es igual a la que existe entre "ali-

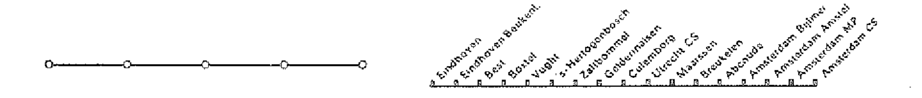


Figura 10. Cadenas.

mentación" y "descanso", y que ambas son menores que la diferencia entre "alimentación" y "rendimiento", por ejemplo.

No siempre los agrupamientos están motivados por la necesidad de considerar las distancias desiguales entre los valores de una variable. Una motivación ya mencionada es la de incrementar la eficiencia de la codificación, en especial si la cantidad de categorías es grande. Al concebir esa variable como una escala nominal, se pasa por alto el efecto ordenador del agrupamiento.

Cadenas

Las cadenas son series de valores totalmente ordenadas. Una cadena tiene dos valores extremos o finales, cada uno de los cuales posee exactamente un solo vecino inmediato, mientras que todos los restantes tienen exactamente dos. En la figura 10 se suministran dos ejemplos; otros serían la trayectoria (lineal) del tiempo, una lista de espera, todas las escalas de medición y un recorrido sin circuitos cerrados. Las cadenas son las formas predilectas de establecer las variables en las ciencias de la conducta; las diferencias entre las escalas nominales, las escalas de intervalos y las escalas de razones o semilogarítmicas son de métrica.

Circuitos cerrados

Los circuitos cerrados [*loops*] son cadenas sin extremos y en las que cada valor tiene exactamente dos vecinos inmediatos, sin que ninguno sobresalga particularmente respecto de los demás. De hecho, si a partir de un valor cualquiera uno se mueve en una dirección, a la larga volverá al punto de partida, que de todas maneras es arbitrario. Se muestran algunos ejemplos gráficos en la figura 11.

Entre los ejemplos de fenómenos registrados mediante circuitos cerrados se incluyen la conducta cíclica o repetitiva, las fluctua-

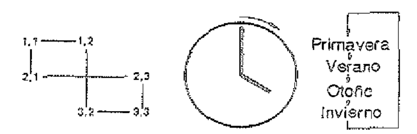


Figura 11. Circuitos cerrados.

ciones estacionales, las preferencias incongruentes y las autorreferencias. En muchas culturas se adoptan concepciones circulares del tiempo; los biólogos acostumbran a describir los fenómenos que estudian en función de los ciclos vitales, y los especialistas en cibernética tienden a rastrear los efectos de los circuitos de autoalimentación en los sistemas complejos. En una red ferroviaria, un recorrido que no implique pasar dos veces por el mismo tramo de vía constituye un circuito cerrado. Otro ejemplo de circuito cerrado bidimensional es el sistema de coordenadas de la superficie terrestre. Aunque las técnicas analíticas tradicionales rara vez pueden abordar los circuitos cerrados, y lo que hacen más bien es abrirlos y convertirlos en una cadena, lo cierto es que estos circuitos poseen importantes propiedades que no pueden representarse mediante cadenas.

Cubos

Los "cubos" incluyen representaciones pluridimensionales de los datos, en las que cada valor tiene tantos vecinos inmediatos como dimensiones existen, y la cantidad de trayectorias paralelas entre dos valores cualesquiera depende de su disimilitud o de la

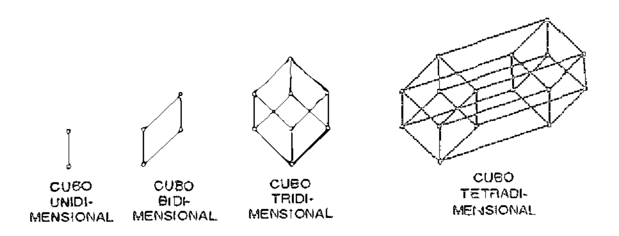


Figura 12. Cubos de una a cuatro dimensiones.

distancia que los separa en una geometría euclidiana (véase la figura 12).

Los cubos suelen plantearse implícitamente. Considérense, por ejemplo, estas ocho categorías de valores de LASSWELL y KAPLAN (1950):

poder
rectitud
respeto
afecto
riqueza
bienestar
erudición
habilidad

Superficialmente, estos ocho valores se asemejan a una escala nominal en la que no existe ningún orden aparente; pero Lasswell y Kaplan establecen que toda unidad de registro, persona, símbolo o expresión, puede tener una puntuación alta en más de un valor. No obstante, esto transgrediría el requisito de que los valores sean mutuamente excluyentes, requisito propio de su consideración como una variable no ordenada. No se trata, pues, de una escala nominal. De hecho, estas categorías de valores son atributos, y toda consigna de "codificar tantos como corresponda" define un cubo cuyos valores son combinaciones (una selección dentro de un número fijo) de atributos.

Arboles

Los árboles o jerarquías ramificadas tienen un valor nominal determinado, por un lado, y varios valores terminales por el otro; y según cual sea el lado a partir del cual se construya el árbol,

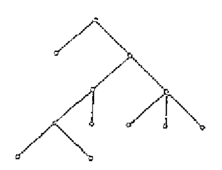


Figura 13.

todos los restantes valores ocuparán o bien puntos de bifurcación o bien puntos de confluencia, respectivamente. Los valores terminales de los dos lados están comunicados entre sí por cadenas que pueden tener algunos miembros en común. Un ejemplo gráfico de un árbol se representa en la figura 13.

Los *nombres* del sistema de Linneo para la clasificación de las especies biológicas constituyen un árbol. El sistema parte de la concepción más general de los seres vivientes en su extremo superior, en un punto intermedio diferencia los mamíferos de los reptiles, y termina con la más pormenorizada diferenciación entre las variedades o especies y cultivos del mundo biológico. Otros ejemplos serían un árbol genealógico, un organigrama de decisiones o los procesos ramificados que suministran trayectorias optativas en cada situación, etcétera.

Los árboles poseen considerable importancia en el análisis de contenido, ya que son capaces de representar distintos niveles de abstracción que son comunes en las expresiones en lenguaje natural. De acuerdo con la teoría aristotélica del significado, una definición nombra el género a que pertenece el *definiens* (la palabra que va a ser definida) y lo distingue de todas las restantes especies de ese género. Una referencia a Europa alude implícitamente a Francia, Italia, Alemania, etc. La relación entre "Europa" y "Francia" es de una implicación unidireccional (la extensión de este último concepto incluye la extensión del primero). En un árbol se representan de hecho estas implicaciones. Las posiciones o cargos en el interior de una organización social jerárquica también forman estructuras ramificadas.

Es importante no confundir los árboles con los agrupamientos, por un lado, y con los reticulados de partición, por el otro. En lo referente a la primera confusión, el sistema de Linneo agrupa sin lugar a dudas un conjunto de seres vivientes que, por lo demás, no tiene orden alguno. No existe ningún animal abstracto llamado "mamífero", por ejemplo. "Mamífero" es el nombre que se asigna a una clase particular de organismos. Los nombres pueden poseer diversos grados de generalidad, no así los animales. El sistema de Linneo define en rigor un árbol de nombres, mientras que agrupa a los animales así nombrados. En lo que respecta a la segunda confusión, digamos que Europa puede sin duda alguna dividirse en países, cada uno de los cuales a su vez puede dividirse en regiones, distritos, etc. Mientras que un árbol puede considerarse como una cadena en un reticulado de particiones, los valores del

árbol representan las partes de dichas particiones pero no las particiones en sí mismas.

Reticulados de partición

Los reticulados de partición son variables que tienen como valores sus particiones o divisiones. Una partición es un conjunto formado por los conjuntos de elementos mutuamente excluyentes. En todo reticulado de partición, dos particiones cualesquiera tienen más esquinas (en las que están representadas las distinciones entre una y otra partición) y una articulación (que representa las distinciones que ambas particiones tienen en común). Esto se muestra en la figura 14:

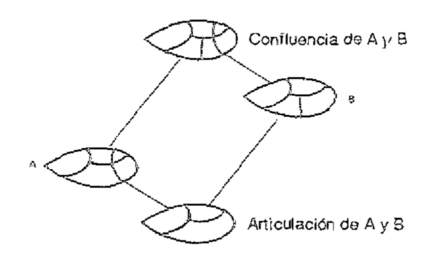


Figura 14. Confluencia y articulación de dos particiones.

Los reticulados de partición son más corrientes de lo que podría pensarse en la vida cotidiana, aunque rara vez se incorporan al catálogo de los procedimientos descriptivos de las ciencias sociales. Considérese, por ejemplo, la sugerencia de LASSWELL (1963) en el sentido de que la política se ocupa de estudiar "quién obtiene qué, en qué momento y de qué manera". Según esto, la ciencia política desarrolla teorías acerca de la forma en que los recursos se distribuyen entre los agentes políticos, y reúne datos al respecto. Presumiblemente, los intereses políticos se manifiestan en las particiones que cada cual tiene presente, y el proceso político procura reducir las diferencias entre la superposición universal y la articulación universal de las particiones. Si se tiene en cuenta la afirmación de Lasswell, gran parte de la teoría política exigiría que los datos se recogieran en forma de reticulados de partición. Estos son indispensables en las investigaciones sobre formación

de coaliciones y en otros procesos semejantes, que dan como resultado distintas divisiones de una misma totalidad en partes mutuamente excluyentes.

METRICA

Dentro de las variables ordenadas, las diferencias entre valores pueden poseer distintas propiedades, que constituyen lo que se denomina una "métrica". Una métrica se define por la clase de operaciones matemáticas aplicables a las variables que no deforman las diferencias representadas en ellas ni introducen cantidades falsas. Por ejemplo, las operaciones algebraicas de suma y multiplicación pueden tener sentido al registrar cifras de población, pero no cuando simplemente se ordenan por su preeminencia los rasgos de determinados personajes. En la bibliografía se distingue entre métrica ordinal, métrica de intervalos y métrica de razones (STEVENS, 1946), aunque no siempre se hace con estos nombres, mientras que las variables no ordenadas o las escalas nominales se definen por la ausencia de toda métrica.

Métrica ordinal

La métrica ordinal proviene de efectuar, entre las unidades de registro, comparaciones del tipo de "mayor que", "más que", "antecede a", "causa", "es una condición de", "es un producto mejorado de", "se transmite a", "está incluido en", "controla", y muchas más por el estilo. Probablemente las disciplinas donde más común resulta el uso de escalas ordinales (cadenas con una métrica ordinal) sean las ciencias sociales. Considérese la siguiente codificación de un índice acerca de la importancia de los artículos periodísticos expresada en función de su diagramación tipográfica:

1) Ocupa el cuadrante superior izquierdo de la primera página.
2) Ocupa el cuadrante superior derecho o el cuadrante inferior izquierdo y el centro de la primera página.
3) Ocupa el cuadrante inferior derecho de la primera página.
4) Ocupa la parte superior al pliegue central de una página interior.
5) Ocupa la parte inferior al pliegue central de una página interior.

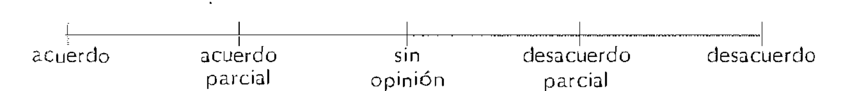
Los juicios relacionados con la preeminencia relativa de un atributo pueden registrarse así:

ausente---------------- apenas determinable ------------- presente pero no predominante --------- predominante

Veamos ahora cómo se construye una escala de identificación para personajes televisivos:

(++) héroe absoluto
(0+) buen tipo / buena chica
(00) a veces bueno y a veces malo, o indefinido
(0-) mal tipo / mala chica
(--) malo absoluto

Aun cuando gráficos como el siguiente



parecen sugerir la presencia de distancias equivalentes, sin contar con pruebas empíricas que avalen el grado en que estos valores están separados uno del otro, todo lo que puede asegurarse es su rango. Una métrica ordinal convertiría en estructuralmente equivalentes estas tres cadenas.

acuerdo	>	acuerdo parcial	>	sin opinión	>	desacuerdo parcial	>	desacuerdo
5	>	4	>	3	>	2	>	1
2570	>	569	>	568	>	12	>	0,001

Métrica de intervalos

Las métricas de intervalos pueden representar ciertas diferencias cuantitativas entre las unidades de registro, concretamente las que pueden expresarse en términos de distancia, similitudes o asociaciones. Entre los ejemplos típicos pueden mencionarse las distancias recorridas, el tiempo transcurrido, el balance o saldo fluctuante (de dinero), los movimientos de actitudes o creencias dentro de un espacio adecuado, las medidas de las diferencias en

la atracción interpersonal, el grado de acuerdo, el grado de apoyo político o la cantidad de comunicación, todo lo cual implica alguna noción cuantitativa de "hasta qué punto son diferentes" dos valores dados. Una métrica de intervalos es la métrica de un espacio continuo, que puede tener numerosas dimensiones, pero que no tiene principio ni final. La representación finita de un espacio de este tipo no es más que una sección arbitraria de dicho espacio extraída del continuo gracias a una elección efectuada por el analista.

Las escalas de intervalos han sido tradicionalmente la columna vertebral de las investigaciones empíricas en las ciencias sociales. El análisis de variancia, las técnicas de correlación y el análisis factorial, por ejemplo, requieren cadenas con métricas de intervalos. En el análisis de contenido estas escalas no son tan frecuentes.

Métrica de razones

La métrica de razones o de cocientes incluye un punto que es el cero absoluto en relación con el cual se expresan todos los valores. La longitud, el peso, la velocidad, la temperatura absoluta (en grados Kelvin), el nivel de ingresos, los centímetros de columna en una publicación, las frecuencias, pero también la cantidad de desacuerdo y de comunicación, son ejemplos de escalas de razones cuyo punto cero es absoluto (a diferencia de los grados Fahrenheit o de los grados centígrados para medir la temperatura, que pueden ser negativos). Otro ejemplo es la escala polar con dos puntos de referencia opuestos, como la usada para enunciar juicios entre los polos de lo bueno y lo malo, entre el cero por ciento y el ciento por ciento, entre lo que predomina o lo que está ausente, en cuyo caso los valores se cuantifican con relación a ambos extremos. Estas clases de escalas se calculan a menudo como escalas de intervalo por el único motivo de que no se dispone fácilmente de las técnicas de análisis apropiadas. Una métrica de razones es la métrica de un espacio continuo para un número cualquiera de dimensiones, con uno (o dos) punto(s) de referencia, a partir del cual (o de los cuales) se considera que los vectores tienen su origen o su destino (o ambas cosas).

Operaciones matemáticas

Como ya se ha indicado, las métricas pueden definirse en función de la clase de operaciones matemáticas que no afecten la diferencia entre las unidades de registro representadas por los valores de una variable. Por ejemplo, si simplemente se añade "4" a todos los puntos de una escala de intervalos que va de "-3" a "+3", lo que se obtiene es una escala cuyos valores van ahora de "1" a "7", pero en la cual las diferencias numéricas siguen siendo las mismas. Cuando los valores de una escala de intervalos se multiplican por una constante, las diferencias numéricas cambian, pero su ordenamiento es el mismo que antes.

Las cuatro métricas a que hemos aludido no sólo son distintas entre sí, sino que su eficacia relativa se considera ordenada jerárquicamente. A partir de la métrica de razones, que es la más poderosa, puede derivarse una métrica de intervalos, y a partir de ésta una métrica ordinal, y a partir de ésta una variable no ordenada. En este proceso, la información que contenía el original respecto de las relaciones se va degradando paso a paso, hasta que sólo permanecen las distinciones.

Es importante comprender las propiedades de las distintas métricas no sólo porque limitan la representación de las diferencias reales, sino además porque algunas de las exigencias que plantean las técnicas analíticas se expresan en esos términos. Por ejemplo, todas las formas de análisis basadas en la variancia implican el cálculo de sumas, promedios y diferencias entre los valores. Para que las diferencias sean significativas, es necesaria como mínimo una métrica de intervalos; para que lo sean las sumas y promedios, es preciso que haya transitividad. Esto excluye los circuitos cerrados y exige utilizar cadenas. Así, pues, las formas de análisis basadas en la variancia requieren escalas de intervalos. Una métrica de razones contendrá más información que la que puede abordar una escala de intervalos, mientras que las variables no ordenadas y las que se ajustan a una métrica ordinal contendrán demasiada poca información como para que pueda aplicarse una escala de intervalos. Cada técnica analítica plantea sus propias exigencias en cuanto a la naturaleza del lenguaje de datos.

métrica	Relación R_{xy}	Relación que preserva f ( ) $f(R_{xy}) = R_{f(x)f(y)}$	Orden que preserva f ( ) $f(R_{xy}) > f(R_{xz}) \iff R_{f(x)f(y)} > R_{f(w)f(z)}$
ausente	distinción $x \neq y$	cualquier 1:1 f ( ), permutación	cualquier 1:1 f( ), permutación
ordinal	rango $x > y$	$x' = x + a$	cualquierf( ) que se incremente monótonamente
intervalo	distancia $[x - y]$	$x' - ax$	cualquierf( ) lineal: $x' = ax + b$
razón	proporción $\frac{x}{y}$	$x' = x^a$	cualquier f( ) exponencial: $x' = bx^a$

9. Construcciones analíticas para la inferencia

Después de haber establecido las distinciones anteriores entre los sistemas, normas, índices, representaciones lingüísticas, comunicaciones y procesos institucionales, en este capítulo se ilustran algunos usos corrientes en la operacionalización de las interdependencias entre los datos y el contexto.

En los capítulos anteriores se estableció que la sensibilidad con respecto al contexto era la característica más importante del análisis de contenido. Dicha sensibilidad se pone en juego: 1) siempre que el investigador entienda que el procesamiento de sus datos no debe constituir un obstáculo para sus cualidades simbólicas, y 2) con el fin de asegurar que se mantengan dichas cualidades, los procedimientos analíticos utilizados deben representar características significativas del contexto dentro del cual se consideran los datos. Una construcción analítica operacionaliza lo que el analista conoce acerca de las interdependencias entre los datos y el contexto.

En su forma más sencilla, una construcción analítica es un conjunto de enunciados del tipo "Si ... entonces ...". Estos enunciados deben contar con una base empírica. En la práctica, la construcción analítica de un análisis de contenido puede parecerse a un modelo de las interdependencias estables en el contexto, siendo

las características inestables aquellas a las cuales pueden apuntar las inferencias. Al igual que los modelos, las construcciones analíticas deben guardar correspondencia, en su función, si no en su estructura, con las características que pretenden representar. El grado de dicha correspondencia mide la sensibilidad respecto del contexto. En la figura 15 se muestra gráficamente esta noción.

También pueden describirse las construcciones analíticas como una teoría acerca de un contexto operacionalizado de modo que sus variables independientes puedan representar todos los datos posibles y sus variables dependientes representen lo que el analista quiere inferir, predecir o averiguar acerca del contexto de sus datos.

Dado que en el momento en que se someten los datos al análisis de contenido no se tiene acceso a su contexto, los conocimientos que se aplican en el desarrollo y justificación de las construcciones analíticas deben obtenerse (o haberse obtenido) mediante otras vías. El origen de dicho saber, el modo en que se operacionaliza y las formas que pueden adoptar las construcciones analíticas constituyen el tema de este capítulo.

Fuentes de incertidumbres

Las inferencias nunca ofrecen certidumbres absolutas. Por consiguiente, el analista de contenido debe evaluar lo mejor que pueda las probabilidades con que los datos disponibles pueden condu-

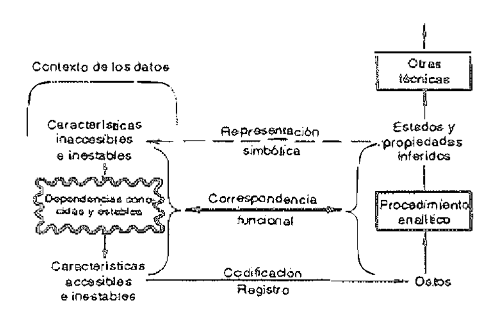


Figura 15. Sensibilidad al contexto en los procedimientos analíticos.

cirlo a las inferencias que pretende extraer. Estas probabilidades provienen de tres fuentes principales:

- Las frecuencias relativas de las dependencias contextuales observadas.
- La confianza en la validez de la construcción analítica.
- El grado de propiedad con que se adecua la construcción a una situación determinada.

Aunque la evaluación numérica de estas propiedades es poco frecuente, es importante examinar qué es lo que influye en estas incertidumbres.

La frecuencia relativa de las dependencias contextuales observadas constituye el caso más obvio. En una argumentación estrictamente deductiva, no hay lugar para las probabilidades: a partir de la premisa "Si A, entonces B", y de datos expresados en la forma de "A", se infiere "B" con certidumbre. Ahora bien: supóngase que las experiencias del pasado se basan en las observaciones siguientes:

B ocurrió después de A	8 veces
B no ocurrió después de A	2 veces
B no ocurrió cuando A estuvo ausente	10 veces

Si estas obervaciones se incorporan de manera adecuada a una construcción analítica (la premisa), entonces, si se observa A, la probabilidad de que ocurra B es 0,8 y la de que no ocurra B es 0,2; mientras que si A está ausente, la probabilidad de B es 0, y su inversa es 1,0. En este sentido, y en las ciencias sociales, la mayor parte de los conocimientos son probabilísticos. Cada regla tiene su excepción.

La confianza en la validez de la construcción es la probabilidad inductiva de que ésta no sea el producto accidental de las circunstancias en que se obtuvieron las experiencias en el pasado. La operacionalización más clara de esta probabilidad se encuentra en la noción de significación estadística. En el ejemplo anterior, un tamaño de muestra de veinte observaciones sería sin duda muy pequeño, y las inferencias que podrían extraerse sobre esa base, sumamente endebles. Si las mismas frecuencias relativas se basaran en doscientas observaciones o más, aumentaría la confianza de que las frecuencias relativas representasen probabilidades auténticas, y el límite superior se alcanzaría cuando la muestra incluyese todos los casos posibles o tuviera un tamaño infinito.

Como no todas las construcciones analíticas proceden de descubrimientos estadísticos, a menudo la confianza en su validez procede de argumentaciones en favor o en contra de la verdad que pretenden transmitir. Para establecer la confianza en la validez de una construcción analítica puede recurrirse a la cantidad de teorías diferentes que conducen a una misma construcción, al número de especialistas que concuerdan con su forma y contenido, y al grado de importancia de las investigaciones emprendidas para recoger las pruebas apropiadas.

El hecho de que una construcción analítica se adecue a una determinada situación parte del reconocimiento de que no existen dos situaciones exactamente iguales y de que las experiencias obtenidas en una condición y un momento determinados, por definidas y válidas que sean, pueden no ser generalizables a otras condiciones o momentos, o sólo serlo en grado limitado. En el análisis del contenido, un buen ejemplo es la creación de una construcción analítica en condiciones experimentales y su aplicación posterior a datos de los que se dispone libremente. En ese caso, la probabilidad de que la construcción conduzca a inferencias válidas depende del grado de realismo que tuviera realmente en las condiciones experimentales. Otro ejemplo es cuando se intenta un análisis de contenido aplicándolo a una situación que se ha modificado subrepticiamente en relación con aquella en la cual se obtuvieron las experiencias respecto de las dependencias contextuales. Este es el problema de los análisis de la propaganda, que suelen ejecutar individuos cuya experiencia respecto del país enemigo se limita a las condiciones anteriores a la guerra. También es el problema que se afronta al analizar documentos históricos con construcciones analíticas contemporáneas. En uno y otro caso, las modalidades de expresión y los significados pueden haber cambiado. Dando por sentado que una construcción analítica se ha desarrollado en condiciones algo distintas de la situación a la que se aplica, la justificación de la propiedad o conveniencia de dicha construcción se reduce a evaluar la similitud entre los datos tomados para el análisis de contenido y los que sirvieron para establecer dicha construcción analítica.

FUENTES DE CERTIDUMBRES

Toda inferencia correcta en un grado superior al azar requiere conocimiento, y si se examinan los argumentos que emplean los analistas de contenido al crear o justificar sus construcciones analíticas, se comprueba que básicamente corresponden a estas cuatro clases:

- los éxitos obtenidos en el pasado
- las experiencias contextuales
- las teorías establecidas
- los intérpretes representativos

Consideremos estas argumentaciones una por una.

Los *éxitos obtenidos* en el pasado aumentan la confianza del investigador en la aplicación de su construcción analítica, siempre y cuando las características que esta última representa sean estables o no se modifiquen en la realidad.

Más allá de este supuesto, el hecho de señalar los éxitos pretéritos no implica conocimiento alguno ni ofrece ninguna intelección particular acerca de la naturaleza de las relaciones contextuales reflejadas por un análisis eficaz.

El empleo del éxito obtenido en el pasado como prueba de una construcción analítica lo ilustran STONE y HUNT (1963) mediante un análisis efectuado por ordenador de notas manuscritas reales y simuladas, elaboradas por suicidas. Las notas reales fueron reunidas a partir de los archivos de los tribunales de Los Angeles; las simuladas fueron redactadas por un equipo de individuos que se basaron en las características de la población (sexo, edad, raza, ocupación, etc.). Se utilizaron quince pares de estas notas, cuya condición real o simulada era bien conocida, para establecer una función discriminatoria o construcción analítica que permitiera al investigador inferir en qué casos una nota era falsa y en qué otros debía tomarse en serio. Se comprobó que las variables discriminatorias eran las siguientes:

1) Referencia a objetos, personas y lugares concretos (mayor para las notas reales).
2) Uso de la palabra "amor" en el texto (mayor para las notas reales).
3) Cantidad total de referencias a procesos de pensamiento y de decisión (mayor para las notas simuladas).

Restando la puntuación de la tercera de estas variables de la suma de la puntuación de las otras dos, se estableció un índice que permitió diferenciar correctamente trece de los quince pares de notas.

Habiendo logrado el éxito en trece de quince casos, la construcción analítica que surgió fue aplicada posteriormente a otros 18 pares de notas cuya condición de reales o ficticias no se reveló a los investigadores. Se verificó que el procedimiento permitía inferir correctamente la autenticidad en 17 de los 18 casos. El hecho de que estas inferencias fueran significativamente más perfectas que los juicios de sujetos obtenidos de forma independiente supuso un apoyo posterior para la validez de la construcción.

En este ejemplo, las tres variables discriminatorias se obtuvieron sin partir de ninguna teoría. Además, la idea de sumar y restar puntuaciones parece ser más bien un artefacto del procedimiento mismo que el efecto de un conocimiento efectivo acerca del comportamiento suicida. Sea como fuere, la función discriminatoria operó convincentemente, y la prueba de este hecho no debe considerarse inferior. En última instancia, todo análisis de contenido debe conducir a inferencias cuya validez sea significativamente mejor que el azar, aunque no sean perfectas.

Las *experiencias contextuales* son a todas luces útiles para la interpretación de los datos, pero son también más personales, menos compartibles, quizá, con otras personas, y por ello tienden a llevar consigo matices de subjetividad. En el análisis de contenido, al igual que en todas las demás actividades científicas, las inferencias tienen que ser independientes del analista, pero los especialistas no pueden apoyarse sólo en su propia reputación. Un analista de contenido que desarrolla una construcción analítica *de novo* debe, como mínimo, explicitar las premisas de sus inferencias. Esta explicitación posibilita compartir dichas experiencias y permite el examen crítico de otros investigadores. Además, facilita el diálogo con la comunidad de científicos sociales, mediante el cual pueden explorarse otras construcciones posibles y surgir argumentaciones contrarias.

LEITES y otros (1951) ofrecen un fascinante ejemplo sobre el grado que alcanzó una construcción analítica "rigurosa" a partir de las experiencias con el contexto del que provenían los datos. Los analistas, todos ellos expertos en la política de la Unión Soviética, estaban examinando la distribución del poder dentro del Kremlin y procuraban formular inferencias acerca de la sucesión

de los líderes soviéticos, proceso que permanece en gran medida oculto para los observadores externos. Contaban con los discursos públicos pronunciados por los miembros del politburó con ocasión del 70º aniversario del nacimiento de Stalin, en 1949; en todos ellos se expresaban las mismas palabras laudatorias hacia Stalin, con matices atribuibles a los estilos de cada cual, y que por ello no revestían interés para el análisis político.

Leites y otros aducían que la clave para diferenciar a los miembros del politburó radicaba en su modo de expresar la "proximidad". Para ello, el uso político del lenguaje que se hace en la Unión Soviética suministraba dos enfoques distintos. Un conjunto de "símbolos de proximidad e intimidad (padre, atención, etc.) aparecen con mayor frecuencia en la imagen popular de Stalin y son realzados frente al público que se encuentra más distanciado de él". El otro conjunto de símbolos procede de "el descrédito en que había caído dicha proximidad en las relaciones políticas. El miembro ideal del partido no pone el acento en las gratificaciones que podría derivar de la intimidad con fines políticos. (...) Los que se hallan políticamente más próximos a Stalin están autorizados a hablar de él empleando términos de menor intimidad personal ('líder del partido', etc.)", y por tanto gozan del privilegio de poder abstenerse de las formas más burdas de adulación. Los autores concluyen sugiriendo que la insistencia relativa de los bolcheviques en la imagen de Stalin (o en su imagen popular) "no sólo refleja la evaluación que hacían los bolcheviques del partido como algo distinto y superior a las masas en general, sino que indica también la distancia relativa del orador respecto de Stalin" (1951, págs. 338-339).

Leites y otros esquematizaron sus descubrimientos y ordenaron a los oradores, comprobando que Molotov, Malenkov y Beria (en ese orden) eran los que reunían la mayor cantidad de referencias a la imagen de Stalin que tenían los bolcheviques, infiriendo de ello que eran probablemente los que más próximos se encontraban a él. La lucha por el poder desatada inmediatamente después de la muerte de Stalin confirmó la plena validez de esta inferencia. Esta construcción fue posteriormente objeto de perfeccionamiento, como se analiza en otro lugar (KRIPPENDORFF, 1967, págs. 118 y sigs.).

Las *teorías establecidas* que relacionan los datos con su contexto son las fuentes más inequívocas de certidumbre para el análisis de contenido. A veces estas teorías toman la forma de propo-

siciones bastante concretas, verificadas en una variedad de contextos (por ejemplo, las relativas a la correlación entre los trastornos del lenguaje y el nivel de angustia del hablante, a la frecuencia de los delitos de que se ha informado y el temor público a una mayor transgresión de la ley y del orden, o al modo en que los televidentes agrupan o clasifican los anuncios publicitarios). En ocasiones, estas proposiciones derivan de teorías más generales, por ejemplo las relacionadas con la expresión de las emociones, con la manifestación lingüística de la psicopatología, o con el modo en que los medios de comunicación de masas y sus respectivos públicos seleccionan y presentan las noticias, y los criterios sobre cuya base lo hacen. Lamentablemente, no existe por ahora ninguna teoría general de la comunicación simbólica, y el especialista en análisis de contenido debe buscar las teorías pertinentes allí donde éstas puedan encontrarse. BERELSON y STEINER (1964) establecieron un inventario de 1025 hallazgos científicos de las ciencias sociales y de la conducta que podrían consultarse.

Un ejemplo del uso de las teorías establecidas para la elaboración de construcciones analíticas se encuentra en OSGOOD y otros (1956) con su "análisis de la aseveración evaluativa". Esta técnica procede de una versión de la teoría de la disonancia, que parte de los siguientes presupuestos:

1) Los conceptos (objetos actitudinales) son evaluados, apreciativamente o no, en grados que van de lo positivo a lo negativo, pasando por lo neutral.
2) Todas las aseveraciones lingüísticas pueden descomponerse en pares de conceptos (objetos actitudinales) que, hasta cierto punto, están asociados o disociados.

y que postula una psico-logía de acuerdo con la cual

3) Los individuos únicamente perciben, expresan o creen en pares de conceptos equilibrados, es decir, en aseveraciones que contienen asociaciones entre conceptos evaluados de manera similar o que contienen disociaciones entre conceptos evaluados de manera diferente. Los pares de conceptos desequilibrados, o sea, las aseveraciones que contienen disociaciones entre conceptos evaluados de manera similar y asociaciones entre conceptos evaluados de manera diferente, son rechazadas o modificadas hasta alcanzar el equilibrio.

Los autores dan un paso más y postulan una relación cuantitativa entre el grado de evaluación positiva (o negativa) de un concepto, el grado de evaluación positiva (o negativa) de un segundo concepto, y el grado de asociación (incluida la disociación) entre ambos según el cual las aseveraciones están equilibradas (o desequilibradas).

La construcción analítica resultante lleva al investigador a inferir actitudes y evaluaciones implícitas a partir de los conceptos explícitamente evaluados y de las asociaciones entre ellos. En un análisis de aseveración evaluativa, se define una estadística de las actitudes explícitas e implícitas que aparecen en un texto, y se llega así a un cuadro global de las actitudes de un autor.

Como es natural, el empleo de una teoría establecida también tiene sus inconvenientes. En el curso de la derivación de proposiciones específicas y de la definición de una construcción analítica en términos operacionales, puede perderse una parte del contenido de la teoría, o quedar demasiado deformada como para ajustarse a los requisitos específicos de un análisis de contenido. En el análisis de aseveración evaluativa, toda inquietud expresada por un autor respecto de la disonancia cognitiva (por ejemplo, el hecho de que se lamente de que su mejor amigo haya hecho algo horrible) no es manipulable. De acuerdo con la construcción, en este caso el amigo del autor tendría que obtener una evaluación negativa, por más que cuente con múltiples formas de resolver el dilema cognitivo. También se plantea si tiene o no sentido descomponer un texto íntegro en pares de conceptos, y si las diversas evaluaciones explícitas e inferidas pueden añadirse de la manera propuesta. De todos modos, en ausencia de experiencias pretéritas, descansar en teorías intersubjetivamente coincidentes es la mejor estrategia que puede seguir un analista de contenido.

Los *intérpretes representativos* suministran una justificación algo incierta de las "premisas" de las inferencias. La incertidumbre procede de las exigencias antagónicas que plantea recurrir a individuos ya sea como observadores científicamente avezados, o como sujetos no contaminados.

En un extremo, el análisis del contenido debe imponer categorías y definiciones a codificadores que son perfectamente representativos, pero cuya respuesta frente al material que sirve de estímulo ya no es similar a la de los sujetos que presuntamente representan. En el otro extremo, el uso de una muestra representa-

tiva de individuos no adiestrados, que actúe como construcción "analítica" aunque sea de manera implícita, puede ofrecer interpretaciones más válidas, pero dado que sus respuestas no están estructuradas, los propósitos del análisis quedan falseados. Entre ambos extremos se encuentra aquella situación en que se permite a los codificadores utilizar sus concepciones espontáneas para dejar de lado las instrucciones que contrarían la intuición, que se oponen en gran medida a la fiabilidad y sólo subrepticiamente promueven la validez. En uno y otro caso, aunque los intérpretes representativos son eficientes y no pueden causar perjuicio alguno, el analista de contenido no puede permanecer tranquilo, precisamente a causa de la representatividad de aquéllos. De las cuatro fuentes de certidumbre para las construcciones analíticas, ésta es la más débil y la menos defendible.

Para sintetizar: la certidumbre que los éxitos obtenidos en el pasado confieren a una construcción analítica es, de los cuatro criterios, el más claro. Todos los análisis de contenido deben tener éxito, en el sentido de que sus inferencias sean válidas, pero ello requiere una prolongada utilización de la misma construcción con datos similares y en contextos parecidos. Los otros tres criterios pueden proporcionar certidumbres en determinados análisis. El uso de la experiencia de un investigador con un contexto puede proporcionarle incertidumbre si se alcanza algún acuerdo intersubjetivo entre todos aquellos que sostienen haber experiementado con el mismo contexto. Ello exige la explicitación y publicación de esas experiencias, así como un debate en el que se alcance el consenso acerca de su grado de generalidad y de aplicabilidad a una situación determinada. Cuando las construcciones analíticas proceden de las teorías establecidas, la certidumbre viene conferida por el estatuto empírico de estas últimas; esto presupone que se haya estudiado bien el campo de interés del analista, específicamente el que aborda las relaciones entre los datos y el contexto. El recurso a codificadores representativos permite obtener certidumbre en ámbitos que suponen interpretaciones simbólicas convencionales de los mensajes, siempre y cuando las circunstancias de su implicación en el proceso de análisis no menoscabe su representatividad. Esto requiere gran cantidad de individuos y constituye el procedimiento menos justificable.

TIPOS DE CONSTRUCCIONES ANALITICAS

Al ocuparnos de los usos y clases de las inferencias, distinguimos entre los sistemas, las normas, los índices y síntomas, las representaciones lingüísticas, las comunicaciones y los procesos institucionales. La manera en que se formulan las construcciones en cada uno de estos ámbitos de aplicación varía enormemente. En esta sección realizaremos algunos comentarios sobre cada uno de ellos, reservando el grueso para los índices, respecto de los cuales es más corriente el uso de construcciones analíticas.

El enfoque del análisis de contenido basado en los *sistemas* fue examinado anteriormente con referencia a la *extrapolación* de tendencias, pautas y diferencias. Las tendencias implican la observación de una o más variables en distintos momentos; la extrapolación de dichas variables a otros períodos exige construcciones analíticas en forma de función recursiva o autocorrelativa.

Las construcciones analíticas destinadas a identificaciones, evaluaciones y verificaciones exigen *normas* para comparar los resultados del análisis de contenido. Ya mencionamos la evaluación del rendimiento de los medios de prensa, el análisis de las psicopatologías, las "desviaciones" periodísticas, y la aplicación de códigos en la industria de los medios de comunicación de masas. La validez de estas construcciones procede de las instituciones que sancionan su uso y que procuran basarse en las inferencias suministradas. En general, las construcciones analíticas toman la forma de un proceso en dos etapas, en una de las cuales se condensan los datos para poder compararlos, mientras que en la otra se aplica la norma para cerciorarse de la existencia de cualquier posible desviación. Ambas son justificadas por la práctica institucional.

Hemos dicho también que los *índices y síntomas* son variables que se correlacionan con lo que presuntamente indican. La forma fundamental de construcción analítica para lo que se da en llamar índices *directos* puede conceptualizarse como un procedimiento de entrada y salida de datos. Y la materialización más eficaz de una construcción de esa índole incorpora una relación o función matemática unívoca. Esto es lo que está en la base, por ejemplo, de la datación de un artefacto antropológico mediante el procedimiento del carbono radioactivo, o en el acoplamiento de un velocímetro de automóvil a la rueda. En las ciencias sociales los índices tienden a ser mucho más inciertos y las correlaciones pueden no ser perfectas.

Además de estar correlacionados con ciertos fenómenos que interesan, los índices deben escogerse de modo que satisfagan dos exigencias adicionales.

En primer lugar, un índice debe ser lo bastante sensible como para *distinguir* entre los fenómenos que interesan. Las diferencias significativas en los fenómenos deben reflejarse en diferencias notables en el índice, y viceversa. De hecho, muchos índices se construyen con el objeto de decidir entre dos fenómenos, estados o propiedades; por ejemplo, el hecho de que un paciente sea o no clasificable como esquizofrénico, o la elección entre dos libros de texto para ver cuál de ambos es más legible, o la comparación de dos programaciones televisivas para ver si la de hoy, por ejemplo, es más violenta que la de ayer. PAISLEY (1964), que pasó revista a todos los trabajos efectuados para inferir los autores de documentos no firmados, enunció diversos criterios con propósitos de diferenciación:

- Un índice debe presentar un pequeño grado de variancia al aplicarlo a toda la obra conocida de un comunicador.
- Un índice debe presentar un alto grado de variancia entre las obras de todos los comunicadores que se comparan.
- La frecuencia que contribuye al valor de un índice debe ser alta con respecto al error de muestreo.

En segundo lugar —y esto es particularmente aplicable a esta forma *directa* de inferencia—, un índice no debe verse afectado por las variables accidentales o irrelevantes respecto del fenómeno indicado. Por ejemplo, MOSTELLER y WALLACE (1963), al tratar de inferir quién fue el autor de los *Federalist Papers*, impugnaron el uso de hombres que hizo YULE (1944), aduciendo que los individuos pueden escribir sobre distintos temas en diferentes situaciones, y en consecuencia la elección de términos de referencia puede contaminar aquello que revelaría la identidad del autor: "Las palabras que queremos utilizar son palabras no contextuales, palabras cuyo índice de empleo sea casi constante por más que se cambie de tema. Por este motivo, nos resultan especialmente atractivas las llamadas palabras funcionales, esas pequeñas palabras que 'llenan vacíos' " (1963, pág. 280).

Una de las peculiaridades del análisis de contenido es el uso de frecuencias como índices directos de fenómenos subyacentes; así se procede al emplear la proporción de palabras que revelan incomodidad como índice de la angustia, o la proporción de espacio

periodístico utilizado como índice del grado de atención. Las construcciones analíticas representan, pues, la correlación existente entre varias de esas frecuencias y varias de las magnitudes que indican estas últimas.

Es importante, pues, señalar que las frecuencias se utilizan aquí de dos maneras diferentes: como índices y como base de correlación entre dos variables. Estos aspectos se confunden a menudo en la bibliografía; por ejemplo, BERELSON (1952) insiste en la cuantificación justificándola principalmente por la necesidad de verificar hipótesis estadísticas (la intensidad de significación de las correlaciones), aunque en sus ejemplos se ocupa primordialmente de las frecuencias como indicadores de otros fenómenos, como la atención y el énfasis.

En apariencia, pasar de contar las frecuencias de palabras, símbolos y referencias, a contar las frecuencias de pares de palabras, co-ocurrencias de símbolos y patrones de referencias propios de los datos, es un tránsito apenas secundario; no obstante, el análisis de contenido tardó cincuenta años en dar ese paso. Si bien BALDWIN (1942) ya había explorado estas ideas en su análisis de las estructuras de personalidad a partir de autobiografías, diez años más tarde POOL (1952) todavía se limitaba a observar que los símbolos tendían a presentarse juntos o en conglomerados, pero sin analizar esta propiedad. OSGOOD (1959) fue quien aplicó estas nociones al análisis de contenido al formularlas como "análisis de contingencia", demostrando su eficacia con respecto al diario íntimo de Goebbels y realizando experimentos para determinar qué indican las co-ocurrencias. Todos los trabajos emprendidos hasta ahora parecen sugerir que, por encima y por debajo del azar, las co-ocurrencias de referencias a determinados conceptos dentro de una corriente de discurso indican asociaciones y disociaciones cognitivas, respectivamente. A partir de entonces las inferencias sobre la estructura cognitiva de los hablantes, autores y receptores del mensaje han encontrado amplia aplicación. El desarrollo y la justificación de las construcciones analíticas que apuntan a estos fines se atienen a la lógica de los indicadores directos de la frecuencia, tal como fueron analizados antes.

Las construcciones analíticas para las formas *indirectas* de inferencia ya no pueden representarse mediante simples artificios *input-output*: reconocen que la correlación sólo está garantizada en condiciones especiales. (Para una elaboración posterior de los métodos indirectos, véase GEORGE, 1959a, 1959b.)

Las construcciones analíticas para las *representaciones lingüísticas* y para las *comunicaciones* son sumamente complejas y superan los límites de este libro. Puede parecer, sin embargo, que siempre abarcan varios componentes, al operacionalizar distintos conocimientos acerca del uso del lenguaje. Uno de esos componentes suministra una descripción sintáctica del material lingüístico que se analiza, otro infiere las posibles funciones lingüísticas y significados de las palabras en su contexto lingüístico, y un tercero traza un mapa de las interpretaciones semánticas incorporándolo a una "concepción del mundo" o "territorio" del discurso, cuya lógica permite al analista formular inferencias acerca de aquello a lo que se hace referencia y aquello que está implícito (HAYS, 1969; KRIPPENDORFF, 1969a). Estas construcciones analíticas no pueden construirse ni validarse sobre la base de correlaciones: se basan en teorías bastante elaboradas, que abarcan procedimientos interactivos y realimentaciones.

Las construcciones analíticas para los *procesos institucionales* no se atienen a ningún formato fácilmente generalizable. Estas construcciones suelen

- ser cualitativas y utilizar modalidades verbales de razonamiento, en lugar de modalidades cuantitativas y formales;
- imponer gran cantidad de restricciones a las posibilidades que están implícitas en las reglas, regulaciones y prácticas conocidas, en lugar de trazar derroteros simples, delineados por leyes;
- utilizar métodos indirectos y múltiples para la formulación de inferencias, en vez de métodos directos y monocausales.

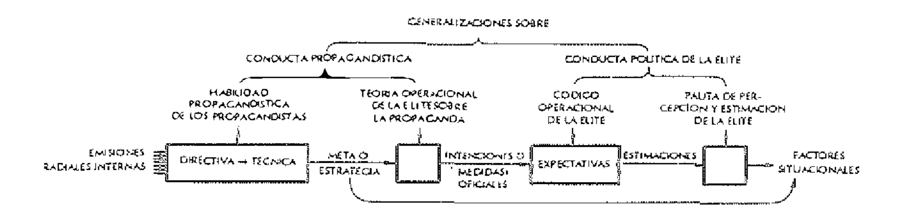


Figura 16. Construcciones analíticas correspondientes a las pautas de comportamiento estable de una élite política.

Aunque la falta de rigor formal de estas construcciones indica incertidumbres respecto de la validez de los resultados obtenidos con ellas, se ha informado sobre el logro de considerables éxitos, en especial cuando se han aplicado repetidamente (permitiendo que exista una validación entre una y otra aplicación) y de una manera prudente (teniendo en cuenta sólo las clases de inferencias que cuentan con mayor grado de corroboración).

Un buen ejemplo es el resumen realizado por GEORGE (1959a) de las inferencias efectuadas por la FCC a partir de las emisiones radiofónicas internas del enemigo durante la segunda guerra mundial. En el curso de su labor, los analistas de la FCC establecieron elaboradas construcciones con el propósito de explicar por qué surgieron determinadas emisiones radiofónicas, y cuáles fueron, por lo tanto, las percepciones y condiciones previas. Las construcciones operacionalizaron: 1) generalizaciones acerca de la habilidad propagandística del principal propagandista; 2) generalizaciones acerca de la teoría operacional sobre la propaganda sustentada por la elite gobernante, es decir, el papel asignado a los medios de comunicación de masas en la prosecución de determinada política; 3) generalizaciones acerca del código operacional de dicha elite, es decir, la manera en que la estimación de la situación y las capacidades se traducían en medidas efectivas; y 4) generalizaciones acerca de la pauta seguida por la elite en la percepción y estimación de sí misma y de su entorno. La figura 16 (tomada de GEORGE, 1959a, pág. 53) muestra de qué manera se concebía que estas generalizaciones regían los nexos causales entre las principales variables inestables del sistema. En este diagrama, las flechas indican el orden en que se formulaban las inferencias; la dirección de la causalidad o la influencia sigue, desde luego, el rumbo opuesto. Probablemente la simplicidad de la figura no refleje la complejidad de la tarea de justificar la sucesión de referencias indirectas que la situación exigía.

10. Técnicas analíticas

En el análisis de contenido, la mayor parte de los esfuerzos en materia de cómputo están destinados a la formulación de los datos y a la aplicación de las construcciones analíticas. En este capítulo se da un paso más allá, hacia el análisis, exploración y descubrimiento de las pautas y relaciones existentes en los datos.

Una vez formuladas las inferencias, o sea, una vez que se conoce lo que significan o indican los datos, se presentan las siguientes necesidades:

- Resumir los datos, representarlos de tal modo que puedan ser mejor comprendidos e interpretados, o relacionados con alguna decisión que el usuario quiera tomar.
- Descubrir en el interior de los datos pautas y relaciones que el "ojo ingenuo" no podría discernir con facilidad, y verificar las hipótesis relacionales.
- Relacionar los datos obtenidos a partir del análisis de contenido con los obtenidos a partir de otros métodos o situaciones, con el fin de convalidar los métodos utilizados o bien de suministrar información ausente.

Estas tareas no son excluyentes; por lo general, se acude a las tres de forma simultánea. Tampoco son exclusivas del análisis de

contenido: gran parte de la labor académica, sobre todo la estadística, tiene que ver con ellas. Dado que prácticamente no existe ninguna técnica analítica de la que el análisis de contenido quiera prescindir, sería imposible presentar aquí una reseña completa de todas ellas. Permítaseme entonces que me centre en unas pocas técnicas que son objeto de una mayor utilización en el análisis de contenido que en otros trabajos empíricos.

FRECUENCIAS

La forma más corriente de representación de los datos es, sin duda, la de la representación de las frecuencias, que desempeña primordialmente la función de compendio del análisis: *frecuencias absolutas*, como el número de incidentes que aparecen en una muestra, o *frecuencias relativas*, como los porcentajes del tamaño muestral. Las *medidas* de cantidad, como los centímetros de columna, el tiempo, el espacio y otros índices basados en frecuencias, tienen en el análisis de contenido el mismo carácter, y no es preciso que las diferenciemos aquí. Salvo en su condición indicativa, como el nivel de atención o el grado en que una actitud o

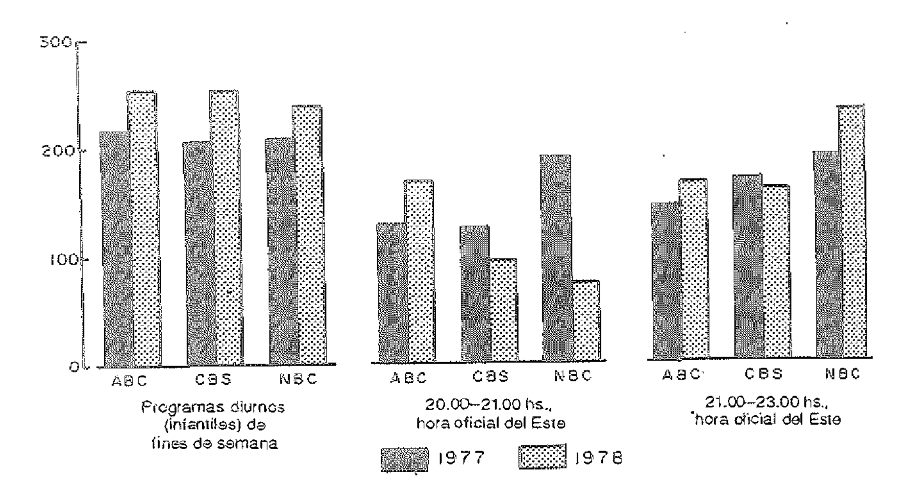


Figura 17. Representación de las frecuencias (índice de violencia) mediante gráficos de barras (índice de violencia). Fuente: *Journal of Communication*, verano de 1979; reimpreso por autorización de George Gerbner y de *Journal of Communication.*

creencia ha penetrado en una determinada población, dichas frecuencias no significan mucho por sí mismas. De hecho, y por lo general, el lector de cualquier estadística suele apelar mentalmente a varias normas para interpretar las frecuencias.

Se apela a la norma de la *distribución uniforme* cuando se descubre que la frecuencia de una categoría es mayor o menor que el promedio de todas las categorías, mientras que se apela a la norma de la *distribución estable* cuando se aprecian modificaciones en las frecuencias con el transcurso del tiempo.

Estas normas suelen estar implícitas en la forma en que se representan o interpretan los resultados analíticos. Por ejemplo, en la figura 17, tomada del trabajo de GERBNER y otros (1979) sobre la violencia televisiva, se apreció que en los programas diurnos de fin de semana la violencia televisiva era mayor que el promedio, y menor que el promedio en los primeros programas de la noche, mientras que en 1977 los programas violentos emitidos por los tres canales televisivos fueron aproximadamente los mismos durante las horas diurnas de los fines de semana. Todo esto apela a la norma de la distribución uniforme. Pero, por otro lado, se observa también que hubo cambios de 1977 a 1978, y si se consideran notables, se apela a la norma de la distribución estable.

Más importante es la norma de la *representación no desviada*, que se invoca cuando se advierte que las frecuencias observadas son mayores o menores de lo que sería previsible si la muestra fuera representativa de la población. No obstante, esta norma no es absoluta, ya que exige el establecimiento de una comparación entre una muestra y alguna otra muestra o población. Por ejemplo, en lo referente a los medios de comunicación de masas, muchos análisis de contenido que utilizaron frecuencias simples advirtieron que la población de los personajes televisivos no es representativa de la población de los miembros del público. Se ha hablado de esto haciendo referencia al grupo étnico, la ocupación, las características socioeconómicas, el sexo y la edad, y suele explicarse mediante los prejuicios, intereses económicos, tendencias tecnológicas y otras construcciones explicativas. Sin embargo, no es obvio que las características del público deban suministrar la norma para los medios de comunicación de masas; quizá sean más fácilmente justificables las referencias a la literatura, el folclore y otras formas de entretenimiento. Generalizando a partir del ejemplo, toda idea de que las proporciones de las frecuencias observadas puedan ser "sorprendentes" requiere hacer explícita la distri-

bución con la que se comparan esas observaciones, y además que la comparación esté realmente justificada.

Asociaciones, correlaciones y tabulaciones cruzadas

Después de las frecuencias, la forma más corriente de representar los datos es mediante las *relaciones entre las variables*. Estas relaciones pueden apreciarse en una tabulación cruzada de las frecuencias de co-ocurrencias de los valores de una variable y de los valores de la otra. Por ejemplo, en un análisis de contenido de actos de violencia televisiva, la muestra, de tamaño n = 2430 actos, dio las siguientes proporciones entre la frecuencia observada/frecuencia *prevista* de co-ocurrencia:

personaje que actúa es	bueno	neutral	malo	
Asociado con el cumplimiento de la ley	369 / *194*	27 / *64*	23 / *161*	419
no asociado con el cumplimiento de la ley	751 / *710*	328 / *233*	454 / *590*	1533
un criminal	5 / *221*	15 / *73*	458 / *184*	478
	1125	370	935	2430

A alguien podría interesarle, no tanto las frecuencias —por ejemplo, el hecho de que la mayoría de los personajes implicados en actos de violencia se presenten como "buenos" (1125) y de ellos, la mayoría como desvinculados de cuestiones legales (751)— sino la relación que podría existir entre el vínculo del personaje con la ley y su presentación favorable o desfavorable. En las frecuencias observadas, esta relación no resulta tan obvia, y es preciso aplicar la norma del *azar o independencia estadística*. Los pares de frecuencias observadas y de frecuencias previstas en condiciones de independencia estadística (que aparecen en el cuadro debajo de las frecuencias observadas) muestran que cualquier informe sobre frecuencias absolutas puede carecer por completo de sentido. La mayor frecuencia observada (751) es también casi igual a la prevista (710) y no contribuye en nada a una posible

relación entre ambas variables. En esta tabla, los casilleros más significativos son los de las cuatro esquinas. Las pruebas de las asociaciones son estadísticamente significativas, más allá de toda duda, lo cual viene en apoyo de la afirmación de que probablemente los "buenos tipos" estén representados como sujetos que respetan la ley, mientras que los "malos" no lo estén.

Las tabulaciones cruzadas no se limitan a dos o tres dimensiones, aunque de este modo es más clara la visualización y más fácil la interpretación. Se dispone de técnicas multivariadas para verificar estructuras complejas con datos multidimensionales (REYNOLDS, 1977).

Las relaciones (asociaciones y correlaciones) pueden establecerse de dos maneras:

- *entre los resultados* de un análisis de contenido;
- *entre dichos resultados y datos obtenidos de forma independiente.*

Dado que el especialista en análisis de contenido goza de pleno control sobre la definición y elección de sus variables, existe siempre el peligro de que las asociaciones establecidas entre los resultados del análisis de contenido sean un producto artificial del instrumento de registro. En el ejemplo anterior, aunque las frecuencias resultaron en cierto sentido sorprendentes, ya que la asociación podría haber sido la opuesta (policías malos, criminales buenos), una correlación positiva, por ejemplo, entre la variable "rasgos de personalidad masculinos-femeninos" y la variable "sexo del personaje" no sería tan sorprendente, ya que ambas variables son *lógicamente codependientes*. A veces, lo único que nos dicen las pruebas tradicionales de la significación estadística o las relaciones entre los resultados de un análisis de contenido es de qué manera definió su instrumento el investigador. En estas condiciones, toda prueba, para ser válida, debe partir de expectativas distintas de las que corresponden a la independencia estadística. En este caso los coeficientes de correlación no se pueden ya interpretar.

Este problema no afecta a las relaciones entre los resultados de un análisis de contenido y otros datos, en gran medida porque estos otros datos se obtienen de forma independiente, por otras vías, en momentos y situaciones distintos. Las pruebas de esas relaciones son sumamente importantes

- para validar los resultados de un análisis de contenido (véase el capítulo 13, sobre la validez), aportando diferentes pruebas al respecto;
- para validar los resultados de técnicas diversas, proporcionando una fuente independiente de pruebas similares (operacionalismo múltiple);
- para verificar teorías que abarcan variables evaluadas por técnicas diferentes.

Ejemplos de esto último son el uso del análisis de contenido para evaluar las respuestas libres a preguntas abiertas en un cuestionario estructurado en todos los demás aspectos; para evaluar los intercambios verbales en el curso de un experimento grupal que, por lo demás, es objeto de una medición minuciosa; y para evaluar el contenido de los medios de comunicación de masas a los que se ha expuesto a determinados miembros del público. La posibilidad de correlacionar la información obtenida a partir del análisis de contenido con los datos procedentes de otras técnicas hace que el análisis de contenido forme parte del amplio sistema metodológico de las ciencias sociales.

IMAGENES, RETRATOS REPRESENTATIVOS, ANALISIS DISCRIMINANTE

Numerosos análisis de contenido se centran en una entidad, persona, idea o acontecimiento especiales, procurando averiguar de qué manera se describe o conceptualiza, o cuál es su imagen simbólica. Como ejemplos de estos estudios pueden mencionarse los referentes a "la imagen del maestro", "la forma en que se refleja el sexo femenino en las películas cinematográficas dirigidas por mujeres", "la manera en que se describe a los Estados Unidos en los periódicos mexicanos" y "lo que el público conoce acerca del Sistema Bell". Existen dos enfoques de este tipo de análisis; son los centrados, respectivamente, en:

- atributos, perfiles de frecuencia, propiedades distributivas;
- asociaciones.

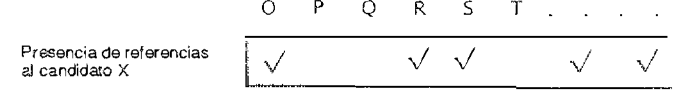
Al estudiar la imagen de un candidato político mediante el enfoque de atribución, el analista de contenido registrará, tabulará

y computará todo lo que corresponde a este candidato, lo que se dice acerca de él, las características que se le atribuyen y quiénes lo hacen, lo que ese candidato ha manifestado, las personas con las que está asociado o sus antecedentes socioeconómicos.

Desde el punto de vista atributivo, la imagen de algo es la representación sistemática de todo lo que según se sabe, o se dice, es propio de ello. Por ejemplo, los documentos relacionados con el candidato en cuestión pueden registrarse teniendo en cuenta la presencia o ausencia de:

O. Rasgos de personalidad favorables.
P. Rasgos de personalidad desfavorables.
Q. Antecedentes como político o funcionario oficial.
R. Actitudes manifestadas respecto de las leyes de asistencia social, etc.

Así, en un documento podemos ver lo siguiente:



	O	P	Q	R	S	T	.	.	.	.
Presencia de referencias al candidato X	√			√	√			√		√

La totalidad de los documentos proporciona un perfil, que consta de las frecuencias de los atributos:

	O	P	Q	R	S	T	.	.	.	.
Frecuencia de las referencias al candidato X	25	10	0	200	186	2	20	.	.	.

Aunque un perfil de esta índole puede representar todo lo que se conoce o se dice acerca del candidato, tiene una limitación: no puede representar más que aquello que se dice acerca de todos los candidatos, lo cual implica que dicho candidato en particular no adquirirá especial realce.

El análisis de la *singularidad* de una imagen o retrato representativo exige hacer comparaciones; en nuestro ejemplo, comparaciones con los demás candidatos políticos. Si se toman todos los atributos correspondientes a la totalidad de los candidatos, y la frecuencia de un atributo cualquiera es igual a la previsible por azar, no puede haber nada singular en cuanto a la relación entre ese atributo y el candidato. Si se desvía de forma sistemática del azar, hay que verificar si esa desviación apunta en una dirección

singular para dicho candidato, o si es compartida con otros. Siguiendo con nuestro ejemplo hipotético, el análisis de imágenes y retratos representativos implica examinar al candidato dentro del contexto de todos los demás candidatos, dividiendo el conjunto de atributos registrados en tres grupos: A, conjunto de los atributos cuyas frecuencias se apartan significativamente del azar en una dirección que no se aprecia en ningún otro candidato; C, conjunto de los atributos cuya frecuencia no se aparta significativamente del azar; y queda entonces B, conjunto de los atributos cuya frecuencia se aparta significativamente del azar, pero en una dirección compartida con algunos otros candidatos.

	A				B					C	
Frecuencia de las referencias	Q	I	·	R	·	·	S	P	·	·	O
al candidato X	0	2	·	200	·	·	186	10	·	·	25
a otros candidatos	113	217	·	11	·	·	78	1	·	·	29
	72	305	·	2	·	·	179	0	·	·	31
	187	212	·	31	·	·	190	18	·	·	19
	93	178	·	17	·	·	207	5	·	·	22
	·	·	·	·	·	·	·	·	·	·	

Los atributos que figuran en A son llamados también los "discriminadores" del candidato, y el análisis que hemos descrito, en su forma básica, es denominado "análisis discriminante".

Desde un punto de vista asociativo, una imagen consta de todo aquello con lo cual está asociada, mientras que excluye todo aquello de lo cual está disociada. El concepto de asociación/disociación es un concepto estadístico que evalúa el grado en que se presentan conceptos seleccionados. La imagen de un candidato pasa a ser así un conglomerado asociativo del cual dicho sujeto forma parte, y que contrasta con los conglomerados de otros individuos.

CONTINGENCIAS, ANALISIS DE CONTINGENCIA

El análisis de contingencia se dedica a inferir la red de asociaciones de una fuente a partir de la pauta de co-ocurrencia de símbolos en los mensajes. Parte de la base de que los símbolos, conceptos o ideas entre los que existe una estrecha asociación conceptual también deben estar estrechamente relacionados desde el punto de vista estadístico. Esto se supone independientemente

de que la fuente sea un autor individual, un grupo social con sus prejuicios o compromisos ideológicos, o toda una cultura con sus pautas de convenciones. Los experimentos demostraron, además, que las asociaciones se transmiten a través de contingencias estadísticas en los mensajes, de modo que el análisis de contingencia puede utilizarse también para formular inferencias acerca de las asociaciones del público. Más allá de estas posibles inferencias, el análisis de contingencia es una técnica analítica por derecho propio.

Parte de un conjunto de unidades de registro, cada una de las cuales se caracteriza por una serie de atributos que están presentes o ausentes. La elección de las unidades de registro es importante, en la medida en que deben ser lo suficientemente ricas, desde el punto de vista de la información, como para contener co-ocurrencias. Una palabra es una unidad demasiado pequeña; una oración contiene por lo general varios conceptos, pero suelen preferirse unidades mayores.

En un segundo paso, se computan y se incorporan como proporciones las posibles co-ocurrencias de los atributos en cada unidad.

En una tercera etapa, debe verificarse la significación estadística de estas co-ocurrencias. OSGOOD (1959), quien formalizó esta clase de análisis, ilustró los resultados mediante un análisis de contingencia de 38 alocuciones pronunciadas por W. J. Cameron en el programa "Ford Sunday Evening Hour". Cada alocución fue considerada como una unidad de registro, y se señaló la presencia o ausencia de 27 atributos de contenido. La pauta de asociaciones resultante se ilustra en la figura 18 (donde las leyendas que aparecen en cada esfera corresponden a la abreviatura en inglés de los atributos).

CONGLOMERADOS

Cuando el cuadro de co-ocurrencias posibles se vuelve demasiado amplio, puede hacerse difícil conceptualizar los resultados. Por ejemplo, el examen de una matriz construida por algo así como 200 x 200 asociaciones entre conceptos (algo que no es tan raro encontrar en los análisis de contenido), representa una tarea formidable, y es probable que al tratar de descubrir una pauta en este alud de información se pasen por alto relaciones importantes.

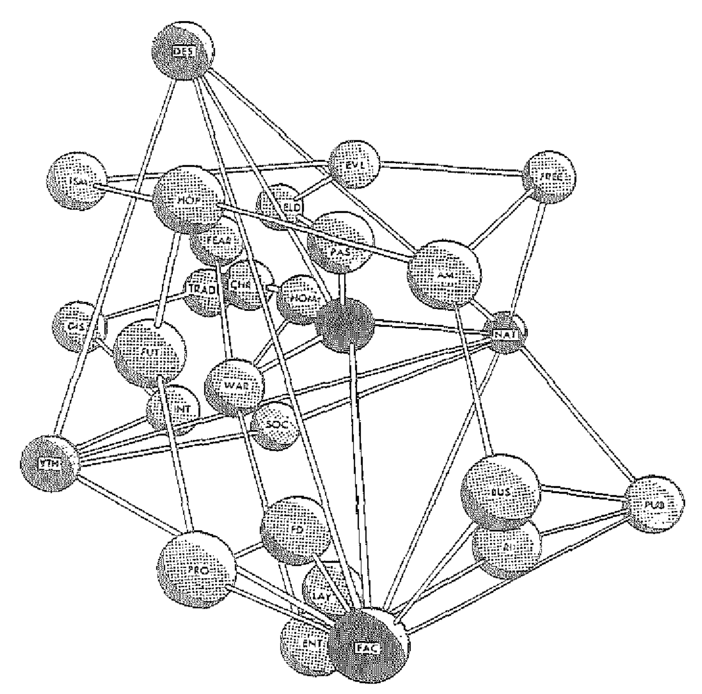


Figura 18. Representación espacial de una estructura de asociación.
Fuente: Ithiel de Sola Pool, comp., *Trends in Content Analysis*,
© University of Illinois Press, 1959.

Afortunadamente, a menudo se comprueba que algunos conceptos son tan similares o están tan interrelacionados, que podrían considerarse uno solo sin perder demasiados matices. Una vez diferenciados muchos "conglomerados" de este tipo, la tarea de conceptualización de los datos se vuelve más sencilla.

La creación de conglomerados procura agrupar o englobar objetos o variables que comparten algunas cualidades observadas, o bien partir o dividir un conjunto de objetos o variables en clases mutuamente excluyentes, cuyos límites o fronteras reflejen las diferencias en las cualidades observadas de sus miembros.

Al formar conglomerados, es importante abordar cuidadosamente el criterio con que se hace. Hay investigadores que prefieren formar conglomerados largos y sinuosos, mientras que otros prefieren los compactos y circulares; algunos son sensibles a la cantidad de detalles que se pierden dentro de cada conglomerado, mientras que otros tienen en cuenta las similitudes entre ellos en

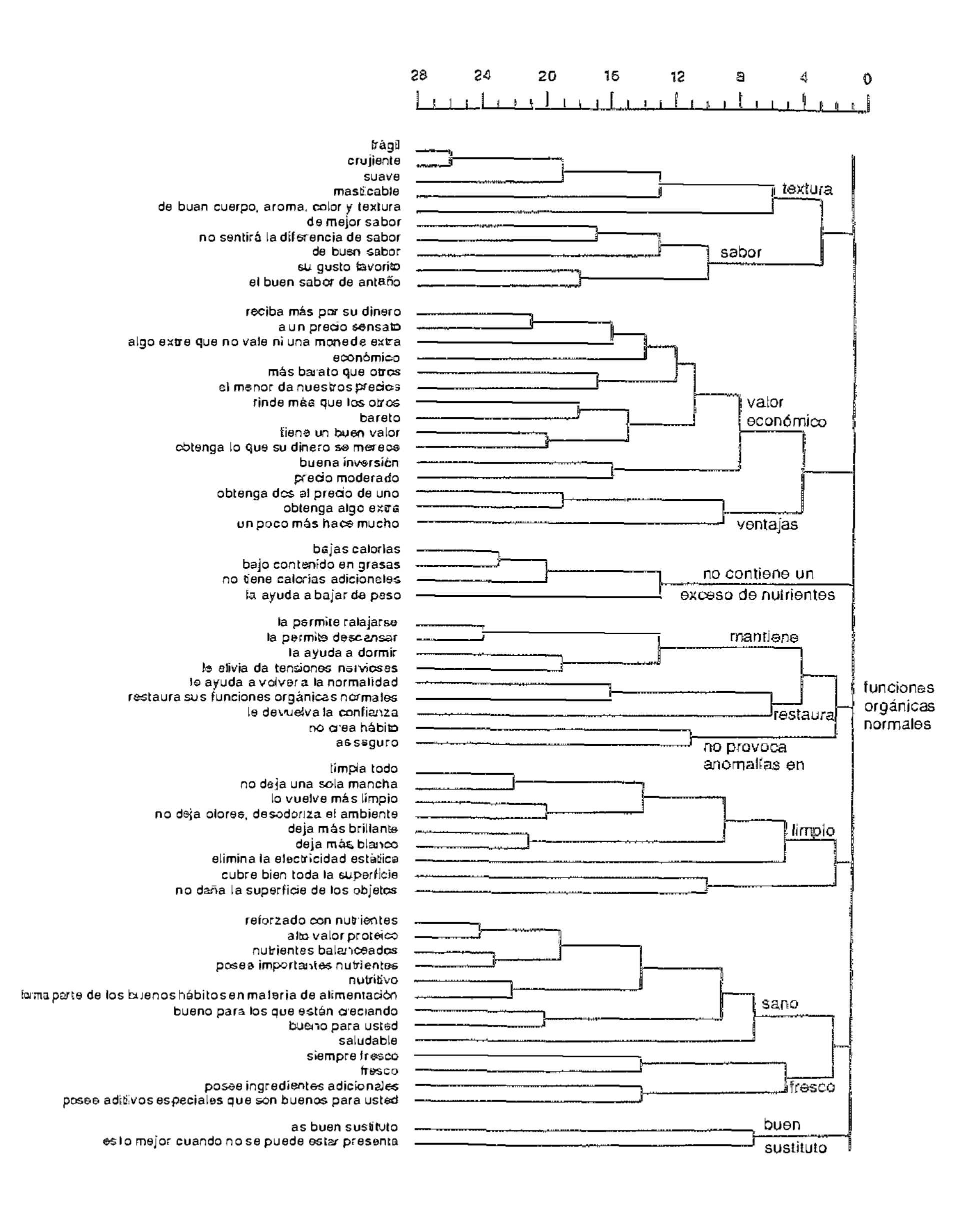


Figura 19. Fragmento de un gran dendrograma resultante de la formación de conglomerados a partir de diversos recursos publicitarios.

su conjunto. Lo ideal es que el criterio de formación del conglomerado refleje la manera en que los conglomerados se forman en la realidad, evaluando las similitudes semánticas más que las puramente analíticas.

Los procedimientos de creación de conglomerados siguen, en esencia, la siguiente trayectoria: en primer lugar, encontrar los dos conglomerados que, según el criterio, son más similares, en el sentido de que su fusión tendrá el mismo efecto sobre las diferencias observadas en los datos en su conjunto. En segundo lugar, agruparlos teniendo en cuenta las pérdidas que se producirán dentro del nuevo conglomerado. Tercero, modificar los datos con el fin de reflejar la última configuración de conglomerados sobre la cual se calculará la próxima fusión. Cuarto, registrar el estado del proceso de formación de conglomerados para el usuario. Los pasos uno a cuatro deben repetirse hasta que ya no quede nada por fusionar. (Para otras referencias al respecto, véase KRIPPENDORFF, 1980.)

Los resultados de un conglomerado pueden presentarse en los denominados "dendrogramas" [por analogía con las "dendritas" de la neuronas], que son diagramas ramificados en los cuales se indica de qué manera los objetos se fusionaron en conglomerados, y en qué nivel de comunidad tuvo lugar la fusión. En la figura 19 se presenta una fracción de un análisis de alrededor de 300 anuncios publicitarios televisivos (DZIURZYNSKI, 1977). La clasificación resultante de estos anuncios en los correspondientes a la textura del producto, su sabor, su valor económico o sus ventajas, tiene considerable validez.

CLASIFICACION CONTEXTUAL

La clasificación contextual es una técnica multivariada para eliminar ciertas clases de redundancia de los datos y así poder extraer de ellos lo que parece ser la conceptualización subyacente. Parte de la base de que los objetos (y, más evidentemente aún, las palabras) tienen más en común (o son tanto más sinónimas) cuanto más se asemeja el contexto en que aparecen. Aquí, por "contexto" se entiende el entorno lingüístico de las palabras o las circunstancias que rodean a una unidad de registro en las proximidades de la fecha en que se produce.

La clasificación contextual comparte algunas exigencias en

materia de datos con la formación de conglomerados, a saber, una cantidad fija de valores para cada unidad de registro. Por ejemplo, en un estudio sobre la percepción de los gobernantes de China, se estandarizó una parte de un discurso de la esposa de Mao Tse Tung (pronunciado el 6 de septiembre de 1968), en el formato fijo siguiente:

1	Srta. Chiang	etapa inicial de la revolución cultural	guardias rojos	enormes contribuciones	al comunismo
2	Srta. Chiang	etapa inicial de la revolución cultural	guardias rojos	enormes contribuciones	a China
3	Srta. Chiang	etapa intermedia de la revolución cultural	guardias rojos	enormes contribuciones	a China
4	Srta. Chiang	etapa inicial de la revolución cultural	guardias rojos	destruyen	a los dirigentes revisionistas
5	Srta. Chiang	etapa intermedia de la revolución cultural	guardias rojos	destruyen	a los dirigentes revisionistas
6	Srta. Chiang	etapa intermedia de la revolución cultural	guardias rojos	aniquilan	la vieja estructura
7	Srta. Chiang	etapa final de la revolución cultural	clase obrera	enseña	a los guardias rojos
8	Srta. Chiang	etapa final de la revolución cultural	clase obrera	enseña	a los guardias rojos

En este ejemplo, las cinco columnas pueden considerarse cinco dimensiones o variables que distinguen los autores, las situaciones, los actores, las acciones y los objetivos. Cada autor, acción, etc., constituye un valor en su dimensión respectiva, y cada enunciado ocupa uno de los 120 casilleros de este espacio inicial. A partir de ahí, la clasificación contextual examina los contextos de cada valor y verifica si los contextos de valores diferentes son lo bastante similares como para considerar dichos valores sinónimos

desde el punto de vista contextual. En lo referente al procedimiento, esto implica reducir el tamaño del espacio multivariado pero no su dimensionalidad, a la vez que se mantienen las relaciones esenciales entre las dimensiones.

En nuestro ejemplo, los enunciados 1 y 2 sólo difieren en cuanto al objetivo de la acción. "China" y el "comunismo" tienen contextos idénticos en todas las unidades, y por consiguiente puede considerarse que tienen un significado similar para su autor. Al tomarlos como una sola clase, es poco lo que se pierde (si es que se pierde algo), aunque la cantidad de casilleros en el espacio se reduce de 120 a 96.

La distinción entre la "etapa inicial" y la "etapa intermedia" de la revolución cultural es el próximo candidato para la clasificación contextual. En los enunciados 2 y 3, y 4 y 5, estas dos "etapas" aparecen en contextos idénticos, mientras que en el enunciado 6, la predicación "aniquila la vieja estructura" sólo se presenta en la "etapa intermedia"; en este caso, alguna distinción se perdería, de hecho, si se descartara esta diferencia contextual. Si esta pérdida fuese tolerable, el espacio se reduciría de 96 a 64 casilleros.

Si se tiene en cuenta el contexto global del discurso político en China, "destruye a los dirigentes revisionistas" en el enunciado 5 es prácticamente sinónimo de "aniquila la vieja estructura" en el enunciado 6, pero en un contexto tan pequeño puede que esto no resulte evidente. Sea como fuere, los próximos pasos del análisis reunirían las acciones y objetivos de los "guardias rojos", por oposición a los de la "clase obrera", en un espacio de 1 x 2 x 2 x 2 x 2 = 16 casilleros.

Aunque la señorita Chiang apoya la convención por la cual un proceso se describe en tres etapas diferentes, y menciona también diferentes acciones y objetivos durante la revolución cultural, la clasificación contextual revela una esencial dicotomía conceptual en cuanto al momento en que se produjo la contribución de los guardias rojos, y en qué consistió ésta, y lo realizado a partir de entonces por la clase obrera.

El procedimiento es más efectivo cuando las frecuencias de los casilleros, dentro del espacio de datos multidimensionales, son grandes (véase KRIPPENDORFF, 1974, 1980); sus límites están determinados por los esfuerzos que exige en materia de cómputo la cantidad de dimensiones en cuestión.

11. Uso de los ordenadores

Si se efectúa enteramente a mano, el análisis de contenido resulta con frecuencia tedioso. En este capítulo se ofrece un panorama de las virtudes y los defectos de los nuevos trabajos que tienden a utilizar técnicas de análisis de contenido basadas en el ordenador, para aplicarlas a determinados contenidos que pueden leerse por medios mecánicos.

Cuando aparecieron los ordenadores digitales, se les dio una amplia y calurosa acogida en el campo del análisis de contenido, creyendo que supondrían un instrumento fiable, rápido y barato; y en realidad, el uso de los ordenadores revolucionó ciertos aspectos del análisis de contenido, que merecen que se les preste especial atención.

Las características de los ordenadores que aquí importan son las siguientes:

1) Gran cantidad de datos digitales los lee el ordenador en forma secuencial.
2) Las operaciones lógicas o algebraicas que pueden definirse en relación con la representación interna de estos datos son ejecutadas a una alta velocidad.
3) La ejecución de dichas operaciones es especificada por un programa que determina y controla "la conducta" del ordenador, y por ello equivale a una teoría perfecta o a una representación sobre el modo en que opera el ordenador.

4) Los procesos de la computación son deterministas, y por lo tanto perfectamente fiables; en el interior del ordenador no se toleran ambigüedades ni incertidumbres.

Aparecen varias analogías entre la forma en que está construido un ordenador y la actividad del analista de contenido. El ingreso secuencial de símbolos delimitados (al que alude el punto 1 de la enumeración anterior) se asemeja a la lectura de expresiones lingüísticas, que también aparecen encadenadas y son digitales. La función del programa en el interior del ordenador guarda analogía con la de las instrucciones que se le dan al individuo humano con el fin de que las aplique de forma fiable: determinación de las unidades, planes de muestreo, instrucciones de registro, procedimientos analíticos... De hecho, suelen utilizarse como sinónimo de "computación" las expresiones "procesamiento de datos" y "manipulación de símbolos", que a su vez se han empleado para describir ciertos aspectos de la inteligencia humana. Veamos en qué aspecto son válidas estas analogías.

En nuestro capítulo sobre los fundamentos conceptuales del análisis de contenido (capítulo 2), caracterizamos a este último por:

a) su naturaleza intrusiva;
b) su capacidad para tener en cuenta material no estructurado;
c) su sensibilidad respecto del contexto;
d) su posibilidad de analizar gran cantidad de datos.

La forma más evidente de ayuda que los ordenadores pueden ofrecer al analista de contenido consiste en el procesamiento de gran cantidad de datos (punto d) a una alta velocidad (punto 2 anterior). Un buen ejemplo, y quizás el ideal, es el análisis "en línea" (es decir, por conexión directa con el ordenador) de impresos. De Weese (1977) creó un mecanismo para convertir las instrucciones tipográficas (a partir de las cuales se elabora la matriz para imprimir el periódico o revista) en una forma analizable, lo cual le permitió someter a análisis a un periódico mientras se encontraba aún en el proceso de ser impreso. Otro ejemplo es la creación de una lista de concordancias de frases o citas correspondientes a toda la producción literaria de un autor. En el pasado, esta labor habría ocupado fácilmente la vida entera de un erudito, mientras que ahora puede llevarse a cabo en unas pocas horas

(siempre y cuando el texto se encuentre en una forma que resulte legible para el ordenador). Y aun con los nuevos procedimientos para la lectura óptica, se ha informado de avances que permiten vislumbrar la posibilidad de introducir en el ordenador gran cantidad de material escrito sin la intervención de un sujeto humano.

La capacidad del ordenador para procesar material textual de forma fiable (punto 4) también ha resultado atractiva para muchos analistas de contenido. Esta elimina prácticamente el error humano, y a la vez convierte en superfluo el esfuerzo por establecer instrucciones de codificación fiables. Pero manteniendo esta virtud como telón de fondo, con frecuencia se olvida que el ordenador utilizado en el análisis de contenido debe comprender, en cierto sentido, el texto, o al menos retener sus cualidades simbólicas (punto c). Mientras que en los análisis de contenido tradicionales se podía confiar en el sentido común de los codificadores humanos y en su habilidad para interpretar e identificar sutiles diferencias simbólicas, si se quieren reemplazar por ordenadores es preciso desarrollar programas adecuados (punto 3). Estos deben ser explícitos y minuciosos, y representar los procesos simbólicos implicados sin abandonar nada a la intuición. No se trata de saber si la fiabilidad de un análisis de contenido puede mejorarse gracias al ordenador, sino de saber si el investigador conoce lo suficiente acerca del contexto de sus datos y puede explicitar este conocimiento en la forma de un programa de computación.

Superficialmente, parece que la necesidad de ser explícito es la causa de la pretendida virtud del análisis de contenido acerca de aceptar material no estructurado (punto b). Pero tanto el ordenador como los analistas científicos se asemejan en cuanto imponen su propia estructura al texto de entrada; la diferencia es que los codificadores humanos pueden escandalizarse o confundirse si comprueban que las categorías analíticas que se les pide emplear son totalmente ajenas a la naturaleza simbólica del texto, mientras que el ordenador simplemente cumple con su tarea, dejando al analista la de preguntarse por el significado del texto de salida.

Los usos que se le da actualmente al ordenador en el análisis de contenido podrían dividirse en tres especies principales:

- análisis estadísticos
- asistencia por ordenador para estudios y descubrimientos
- análisis de contenido por ordenador

ANALISIS ESTADISTICOS

En las ciencias sociales, las técnicas estadísticas son procedimientos corrientes, y se puede tener acceso prácticamente a todos ellos en la forma de programas de computación. En realidad, casi nadie hace hoy un análisis factorial a mano (y tal vez ya no haya personas en condiciones de hacerlo). Además de los programas individuales, existen grandes "paquetes" de soporte lógico estadístico (SPSS, BMDP, LISREL, DATA-TEXT) de uso sencillo incluso para las personas que no están familiarizadas con la programación de ordenadores.

Para aplicarlos, los datos se codifican con el fin de ajustarlos a las exigencias técnicas en materia de datos de entrada, mientras que las opciones computacionales son especificadas por el usuario. En el curso de la computación, los datos son clasificados, reorganizados, transformados y descritos por índices numéricos, que el usuario debe luego interpretar según lo que haga la técnica. Así, pues, es esencial una cierta comprensión de la forma en que se producen estos índices.

En los casos en que los procedimientos estadísticos se emplean *como medio de formular inferencias*, o cuando se incorporan como parte de un proceso global de formulación de inferencias (véase el capítulo 10, "Técnicas analíticas"), su estructura interna debe ser sensible al contexto (véase el capítulo 9, "Construcciones analíticas para la inferencia"), es decir, deben representar los procesos simbólicos en los que intervienen los datos. En dichas aplicaciones, el usuario tendrá quizá que familiarizarse con la bibliografía teórica sobre esos procedimientos y decidir si las premisas incorporadas a éstos se justifican realmente, en cada contexto particular.

Más habitual en el análisis de contenido es aplicar las pruebas estadísticas *después de haber efectuado las inferencias*, por ejemplo para resumir las clases de símbolos o de referencias efectuadas en un texto, para verificar la independencia estadística de los atributos, o para relacionar los hallazgos del análisis de contenido con los resultados obtenidos mediante otros métodos de investigación. No hay entonces necesidad de justificar estos usos más allá de lo referente a sus propiedades formales.

ASISTENCIA POR ORDENADOR PARA ESTUDIOS Y DESCUBRIMIENTOS

Antes de conceptualizar sus datos y decidir si habrá de recurrir a una técnica de reducción de datos, a la elaboración de un diccionario o a los procedimientos analíticos que podrían adecuarse a su caso, el investigador tal vez desee disponer de un panorama acerca de la variedad, clases y distribución de los datos. La asistencia por ordenador para obtener este panorama está motivada primordialmente por el gran volumen de material textual que el analista debe considerar; carente de esta ayuda, es probable que se cree impresiones tendenciosas, incompletas y muy selectivas.

Una ayuda simple la proporciona la lista alfabética de las unidades de análisis, por ejemplo las palabras que aparecen en un texto. Se ha probado que esto facilita también la corrección de errores tipográficos en el flujo de datos de entrada.

Otra manera, tal vez más instructiva, de alistar las unidades es hacerlo por su frecuencia de aparición. Y esta lista puede reducirse considerablemente si se atiende a las raíces o temas de las palabras o a su clase, en lugar de diferenciarlas por sus variantes específicas. Por ejemplo, si se pasa por alto la desinencia o inflexión, las palabras "hablará", "hablaría", "hablaron", "hablaba", etc., se reducen a una sola. Experiencias efectuadas con uno y otro tipo de lista muestran que tanto la frecuencia de las palabras como la de sus raíces sigue una curva exponencial. Habitualmente hay unas pocas palabras de alta frecuencia (artículos, palabras funcionales que no distinguen a un documento de otro) así como un gran número de palabras que se presentan de forma aislada y que no mantienen asociaciones ni correlaciones significativas con otras. De esta manera, el simple cómputo de las frecuencias de las unidades presentes en sus datos permite al investigador formarse una idea acerca de las cantidades que debe abordar, así como también establecer ciertas decisiones simplificadoras, por ejemplo eliminar del análisis las palabras más frecuentes o las menos frecuentes.

Otra ayuda la ofrecen los programas de recuperación de información. En esencia, estos programas cumplen una serie de instrucciones de "búsqueda y localización", que pueden formularse ya sea sobre la base de una teoría *a priori*, según la cual ciertas unidades, palabras o raíces de palabras son más importantes que otras, o bien después de examinar la lista de frecuencias. El producto de un programa de esta índole mostrará la situación o situaciones de cada

unidad, o bien imprimirá todas las oraciones, párrafos o documentos que la contienen, facultando al analista para examinar sistemáticamente la totalidad de la base de datos desde una perspectiva determinada. Una de las primeras aplicaciones de un sistema de recuperación de la información con fines investigativos fue la atribuida a SEBEOK y ZEPS (1958), quienes realizaron una búsqueda referida a las pautas presentes en las leyendas y cuentos folclóricos escritos en cheremis. JANDA (1969) creó otro sistema semejante, que almacena documentos en microfilms y permite al analista ir añadiendo indicadores de contenido (categorías) a medida que examina el material. En este caso la totalidad del documento constituye la unidad de recuperación. EKMAN y otro (1969) describen un sistema similar para la exhibición de información visual, utilizado en el análisis de expresiones faciales.

Todos los cómputos y listados alfabéticos arrancan a las palabras de su entorno lingüístico original y, por ello, sus significados dependientes del contexto ya no resultan reconocibles por el analista. En cambio, el método conocido como Key-Word-in-Context (KWIC) hace exactamente lo opuesto: enumera las apariciones de palabras seleccionadas junto con el entorno lingüístico en que se presentan, y da así al investigador una idea de la forma en que se utiliza cada palabra clave. En la figura 20 (tomada de STONE y otros, 1966, pág. 159) se reproduce el ejemplo proporcionado por Dunphy de una lista KWIC para la palabra inglesa *play*. Si se conservan las referencias a los documentos originales, como se hizo en este caso, el analista puede además examinar si los distintos sentidos o acepciones de una palabra coinciden en diferentes autores, documentos o situaciones. Las listas KWIC son asimismo una importante ayuda para establecer categorías de significado, como las que se necesitan en la elaboración de diccionarios y la codificación manual.

Otra ayuda computacional que por su naturaleza se aproxima a una técnica estadística de decisión, consiste en identificar una tira impresa para su examen y seleccionar en ella, con vistas a su análisis posterior, las unidades o palabras que muestran alguna asociación con otras palabras, y que en ese sentido son portadoras de información. El texto en su conjunto es explorado por una especie de "ventana" de, digamos, 200 caracteres o 20 palabras de extensión, y la cantidad observada de apariciones se compara con sus expectativas estadísticas, reteniendo aquellas que muestran un desvío inusual respecto del azar (véase IKER, sin fecha).

A VIOLINIST OF THE HIGHEST CALIBRE	PLAY	A CONCERTO. JIMMY HAS BEEN TAKING	5C1	WILLIAMSON UNDERGRAD
IN HIS ROOM, HE WOULD HALF-HEARTEDLY	PLAY	A FEW SCALES, AND THEN WOULD PUT	5C1	WILLIAMSON UNDERGRAD
LEADERSHIP, THAT SCOTT WILL PROBABLY	PLAY	A LARGER ROLE IN THIS AREA, AND	2H4	SHAPIRO SMALL GROUP
BEST BUDDY. ..HOW ABOUT COMING OVER TO	PLAY	A LITTLE BALL... HIS FRIEND	5C1	WILLIAMSON UNDERGRAD
WHILE HE IS YOUNG HE SHOULD LEARN TO	PLAY	A MUSICAL INSTRUMENT, THE VIOLIN.	5C1	WILLIAMSON UNDERGRAD
A GREAT DEAL TO US) COULD CONCEIVABLY	PLAY	A SIGNIFICANT PART. IT SURPRISES ME	2H5	SHAPIRO SMALL GROUP
AN AIR OF DISDAIN. HOW CAN HE POSSIBLY	PLAY	A SOLO IN THE CONCERT TOMORROW	5C1	WILLIAMSON UNDERGRAD
COME TO PLAY A TUNE I HAVE COME TO	PLAY	A SONG. WILL YOU GIVE ME YOUR	0A3	ZINACANTAN DREAMS
MIGUELITO HOW ARE YOU. I HAVE COME TO	PLAY	A TUNE I HAVE COME TO PLAY A SONG.	0A3	ZINACANTAN DREAMS
A VIOLIN, HE WANTED TO LEARN HOW TO	PLAY	A VIOLIN. GO TO THE NEXT ONES HE	5A1	HARTMAN DETROIT GANGS
SEEM, YOU KNOW, ENCOURAGED ENOUGH TO	PLAY	AN INSTRUMENT AND I DO NOT THINK HE	5A1	HARTMAN DETROIT GANGS
PARENTS HAVE DECIDED THAT HE SHOULD	PLAY	AN INSTRUMENT, AND HAVE BOUGHT HIM	5C1	WILLIAMSON UNDERGRAD
WELL GROWN THEY USED TO GO OUTSIDE TO	PLAY	AND HAVE FRESH AIR. THEY WENT ON IN	0J1	KIKUYU
THINGS, THE CHILDREN WOULD ASK HER TO	PLAY	AND THAT IF SHE SEES THAT IT IS	1A1	FAM INTERACTION FRISBIES
BOY. WHEN I WAS DOWN SOUTH, I USED TO	PLAY	AND THEN I WOULD GET MAD AND GO IN	5A2	HARTMAN DETROIT GANGS
HE WANTED SOMETHING, AND SO WE WOULD	PLAY	AND THROW BOXES ALL OVER THE ROOM,	5A2	HARTMAN DETROIT GANGS
I-VE DECIDED TO TAKE YOUR OFFER AND	PLAY	AT THE ..POPS.. TONIGHT. HOW MUCH	5C1	WILLIAMSON UNDERGRAD
YOU SEE I WILL BE BACK I AM GOING TO	PLAY	AT THE HOUSE OF OUR FRIEND MANVEL	0A3	ZINACANTAN DREAMS
AND SCOLD HIM AND TELL HIM HE CANNOT	PLAY	BALL AT ALL TODAY SINCE HE DID NOT	5C1	WILLIAMSON UNDERGRAD
IS A NICE SPRING DAY AND HE WANTED TO	PLAY	BALL INSTEAD. HE IS NOW TRYING TO	5C1	WILLIAMSON UNDERGRAD
CONSOLE HIMSELF BY GOING OUT TO	PLAY	BALL. AND RESIGNEDLY RETURN TO	5F1	H ACH TAT
HE IS IN THE ATTIC, WATCHING SOME BOYS	PLAY	BASEBALL ON THE OPPOSITE HILL. HE	5C1	WILLIAMSON UNDERGRAD
I GO AROUND THERE WITH MY FRIENDS AND	PLAY	BASKETBALL ALL THE TIME. THAT IS	5A2	HARTMAN DETROIT GANGS
AFRICANS IN BASKETBALL, AND I OFTEN	PLAY	BASKETBALL WITH THE LEBANESE.	2C2	PEACE CORPS
WORK AND SO THEN I GOES DOWN THERE AND	PLAY	BASKETBALL, BECAUSE THAT HELP MIND	5A2	HARTMAN DETROIT GANGS
BIT COCKY + LOYAL TO MY FRIENDS + A	PLAY	BOY + LENIENT TOWARD SOCIAL 20191	9C2	WHO AM I
HE GREW A LITTLE BIGGER HE LEARNED TO	PLAY	BY HIMSELF AND WHEN HE GOT BIGGER	5G2	NAVAHO TATS
HAD TO SIT THERE AND LET THE PIANO	PLAY	BY ITSELF. HE WAS IN SCHOOL AND THE	0B2	CHILDREN S DREAMS
IN THE EVENINGS WE TALK, WRITE,	PLAY	CARDS OR SCRABBLE. WE HAVE A NEW	2H1	RAMALLO
(THAT HE/SC HAS ROBBED/S HIS FATHER).	PLAY	EENY MEENY MINEY MOS WE/1 HAVE/3	1G1	MILLS
BE THAT IN SIX OR SEVEN YEARS HE WILL	PLAY	FAIRLY WELL AND BE EXTREMELY	5C1	WILLIAMSON UNDERGRAD
OURS THOU WHO ARE IN MAPANTAL. DO NOT	PLAY	FALSE BUT TURN THOU SOON ASIDE.	0E1	IFUGAO
JUNIOR YEAR + ON THE HOUSE COMMITTEE +	PLAY	FOOTBALL + PLAY TENNIS + ENJOY	9C4	WHO AM I
I SELDOM PLAY I DONT KNOW AND I PLAY I	PLAY	FOR MY OWN NO I CANT NOW I LEARNED	1F1	JAFFE CASE
I WAS ON THE PITCHING MOUND. I	PLAY	FOR THE DETROIT TIGERS. WELL, I	5A1	HARTMAN DETROIT GANGS
TWO TO TWELVE. WE COMB THEIR HAIR,	PLAY	GAMES WITH THEM, SEW CLOTHES FOR	2C5	PEACE CORPS
TO TALK WITH THEM, SHOW AND	PLAY	GAMES, SING AND WATCH THEIR FACES	2K2	RAMALLO
OBSERVE THE RITUAL PROHIBITIONS. THEY	PLAY	GONGS AND DANCE. ON THE EIGHTH DAY	0E1	IFUGAO
WE WILL GET SUGAR CANE TODAY. THEY	PLAY	GONGS FOR THE SUGAR CANE. THEY CUT	0E1	IFUGAO
WENT STRAIGHT TO KWENENCEDAS HOUSE TO	PLAY	HAVOC. A LONG BLAST FROM A HORN WAS	0J1	KIKUYU
MIGUEL HE SAID. NOW BASTARD LET US	PLAY	HE SAID. TAKE OFF YOUR CLOTHES WE	0A3	ZINACANTAN DREAMS
TO HIS WOMANS HOUSE HE WOULD COME TO	PLAY	HIS GUITAR. WELL NO THE OLD MAN	0A1	ZINACANTAN FOLK
EVEN LEARN THE PIANO AND I I I SELDOM	PLAY	I DONT KNOW AND I PLAY I PLAY FOR	1F1	JAFFE CASE
I I I SELDOM PLAY I DONT KNOW AND I	PLAY	I PLAY FOR MY OWN NO I CANT NOW I	1F1	JAFFE CASE
AND SADDLES. THAT WAS OUR GREATEST	PLAY	IN OUR VILLAGE IN BASUTOLAND. THE	5D1	SULMAN SOUTH AFRICAN

Figura 20. Muestra de Key-Word-in-Context: la palabra *play**.
Fuente: Philip J. Stone y otros, *The General Inquirer: A computer Approach to Content Analysis*, MIT Press. pág. 159. Reimpreso con autorización de MIT Press.
* La palabra inglesa *play* desempeña en estos fragmentos de oraciones diversas funciones gramaticales (verbo, sustantivo) y es utilizada en muy distintas acepciones ("tocar" un instrumento musical, una escala, un concierto; "desempeñar" un pequeño papel, una función significativa; "cantar" una canción, una melodía; "jugar" al baloncesto, a los naipes, etc.). [T.]

Como es obvio, estos procedimientos de asistencia por ordenador abarcan un amplio espectro de técnicas, incluyendo en un extremo los listados cualitativos del tipo del KWIC y, en el otro, la selección cuantitativa de las palabras portadoras de mayor información. Estos métodos de ayuda son aplicables a grandes volúmenes de datos y sólo exigen adoptar algunos supuestos mínimos concernientes a las construcciones analíticas subyacentes. La interpretación de sus resultados queda totalmente en manos del analista. Aunque mantienen algunas referencias a las cualidades simbólicas de los datos, no establecen inferencias.

Lamentablemente, no se ha prestado suficiente atención, en el análisis de contenido, al desarrollo de estas formas de ayuda por ordenador para reconocer pautas complejas y para exhibir las diferencias de significados distribuidas en grandes cuerpos textuales o en otros vehículos. Estos procedimientos cobrarán cada vez más

importancia, sobre todo con el auge del procesamiento interactivo de datos.

ANALISIS DE CONTENIDO POR ORDENADOR

Los usos mencionados de los ordenadores puede que no se traduzcan en inferencias del tipo de las que se buscan con el análisis de contenido, y además, de hecho esos usos no son exclusivos del análisis de contenido. Designaremos con la expresión *análisis de contenido por ordenador* únicamente aquellas situaciones en las que un ordenador se programe con el objeto de imitar, modelar, reproducir o representar algún aspecto del contexto social de los datos que procesa, o sea, para aquellos casos en que el proceso de formulación de inferencias a partir del texto original (o de los datos en general) esté informatizado en gran medida o en su totalidad.

Aunque la finalidad del análisis de contenido por ordenador es aceptar todo texto primitivo como dato de entrada, con frecuencia la máquina no es capaz de reconocer teóricamente ciertas distinciones relevantes, que entonces deben serle introducidas mediante *preedición manual*. Por ejemplo, registrar y analizar las pautas fónicas de los seres humanos no resulta difícil a través de procedimientos mecánicos, pero sí lo es en extremo inventar un método para reconocer las palabras en esas pautas. La transcripción humana de la voz humana es, simplemente, más eficiente. En la práctica del análisis de contenido, la preedición de un texto puede significar: 1) repetir el punto final para indicar el fin de una oración; 2) marcar mediante algún símbolo especial el comienzo de un párrafo o el momento en que cambia el hablante; 3) explicitar las abreviaturas; 4) añadir definiciones especiales; 5) sustituir los pronombres por los nombres propios que les corresponden; 6) identificar a los autores, las acciones, los objetivos y las justificaciones; 7) distinguir, por ejemplo, entre la ciudad de Buffalo y el animal "búfalo", etc. También puede ser necesario explicitar los núcleos de significado de expresiones complejas o volver a escribir todo el texto en un formato corriente, empeño éste que puede contradecir todos los argumentos en favor de la eficiencia de un enfoque por ordenador: ocurre que no es nada sencillo programar un ordenador para que comprenda las comunicaciones simbólicas espontáneas. La mayoría de los sistemas operativos se limitan a

una forma muy restringida de comprensión, propia de un discurso en particular o circunscrita a una forma específica de datos de entrada.

Dentro del análisis de contenido por ordenador, pueden distinguirse dos métodos principales:

- los métodos del diccionario y de los tesauros;
- los métodos de la inteligencia artificial.

Hay que añadir que este campo es muy dinámico y tal vez no se haya desarrollado tanto como sería de desear. En una reseña reciente sobre las aplicaciones del ordenador a las ciencias sociales (*American Behavioral Scientist*, volumen 20, nº 3, 1979) apenas se menciona el análisis de contenido, aunque sabemos que se están realizando grandes esfuerzos en este sentido.

Métodos del diccionario y de los tesauros

Todos los métodos de los diccionarios y tesauros hacen hincapié en las palabras individuales o las breves series de caracteres identificables en un texto, que son apartadas de su entorno lingüístico, clasificadas o "tipificadas", y luego contadas. La clasificación o tipificación [*tagging*] de las palabras (o unidades de registro) tiene por objeto reproducir los juicios coincidentes sobre las similitudes semánticas de los pares de palabras o de los sinónimos, y por lo tanto es de suponer que constituye una parte importante de la interpretación semántica de un texto. Varios enfoques diferentes coinciden en esta simple forma de comprensión. Aquí esbozaremos tres sistemas prototípicos.

El *método del tesauro*, en su forma pura —tal como se presenta en un conjunto de programas de computación denominados VIA (SEDELOW, 1967)—, saca partido de la existencia de tesauros de la lengua inglesa como el *Roget's University Thesaurus* o el *Webster's New Dictionary of Synonyms*. En éstos, las palabras se agrupan de acuerdo con los significados que comparten; además, según el esquema del *Roget's*, el significado compartido se designa mediante un título o encabezamiento, al que se concibe más simple, general y abstracto que los términos que abarca. De este modo, el *Roget's International Thesaurus* (1974) agrupa los sinónimos en ocho clases: "Relaciones abstractas", "Espacio", "Física", "Materia", "Sensación", "Intelecto", "Volición" y "Emo-

ción". En un segundo nivel de clasificación, el ítem ""Intelecto" se subdivide en "Facultades y procesos intelectuales", "Estados mentales" y "Comunicación de las ideas". En esta última categoría se reconocen trece subdivisiones, de las cuales "Naturaleza de las ideas comunicadas" tiene a su vez otras nueve, de las cuales "Significado", presenta a su vez otras quince, y el segundo grupo de nombres pertenecientes al ítem "Significado" termina enumerando las siguientes palabras: *Intención, Propósito. Finalidad, Objeto, Objetivo, Designio, Plan* y *Significación.*

SEDELOW (1967) brinda una convincente demostración de la necesidad de tabular las palabras de acuerdo con los encabezamientos de los tesauros. Esta autora analizó dos traducciones al inglés distintas de la *Estrategia militar* de Sokolovsky, comprobando que diferían aproximadamente en unas tres mil palabras: 1599 de ellas sólo aparecían en la traducción publicada por la casa Rand, mientras que 1331 sólo aparecían en la traducción publicada por Praeger. Dado que ambas eran traducciones respetables, habría sido un error concluir de esto que diferían en contenido por ese motivo. La sustitución de las palabras por su categoría de significado eliminó las simples variaciones estilísticas y las palabras idiosincrásicas elegidas, que explicaban esas diferencias superficiales. La finalidad del método del tesauro es ofrecer una representación más simple del significado de un texto, identificar ciertos conceptos básicos, verificar con qué frecuencia aparecen determinadas ideas y comparar distintos documentos entre sí para determinar qué hay de nuevo o de diferente en cada uno de ellos.

Además de las limitaciones inherentes a todos los análisis de palabras aisladas de su contexto, uno de los problemas corrientes en el enfoque de los tesauros es que los términos aparecen enumerados en más de un grupo debido a su similitud, lo cual hace más difícil interpretar las descripciones estadísticas de los encabezamientos. Por ejemplo, el *Roget's International Thesaurus* incluye la palabra *intención [intent]* en nueve grupos diferentes, que a su vez aparecen en cuatro de las ocho clases primarias. El analista tiene que decidir hasta qué nivel de clasificación habrá de simplificar el significado de un término. Un segundo problema (que paradójicamente se considera el punto fuerte de este método) es la ausencia de concepciones teóricas en las categorías del tesauro. Sin embargo, siempre que un análisis de contenido intenta algo más que una interpretación semántica general, completa y compartida de las palabras de un texto, esta ausencia se convierte en

un punto débil. Frente a esto, debe señalarse que aun cuando un tesauro pretende establecer las distinciones de acuerdo con el lenguaje al uso, no resulta en absoluto claro de qué manera ha extraído sus categorías: tal vez el *Roget's*, no esté libre de las tendencias conceptuales propias de un autor del siglo XIX, y omita distinciones que hoy resultan importantes.

El *método del diccionario* para el análisis de contenido encuentra su mejor ilustración en el General Inquirer de STONE y sus colaboradores (1966). Este sistema, a diferencia del método del tesauro, incorpora un esquema clasificatorio de origen teórico para las palabras o raíces de palabras que se presentan en un texto. El desarrollo del General Inquirer fue, con mucho, el empeño más ambicioso por informatizar el análisis de contenido, y ha encontrado desde sus orígenes numerosas aplicaciones, modificaciones y traducciones a otras lenguas distintas del inglés. Describiremos a continuación sus características principales.

El sistema acepta un texto primitivo o preeditado en una forma que sea legible para el ordenador (habitualmente en tarjetas perforadas), transfiriéndolo luego a una cinta. Como primer paso, se identifican las raíces de las palabras suprimiendo los sufijos del tipo de "-s", "-ed", "-ing", "-ion" [que en inglés indican, respectivamente, la tercera persona del singular o bien el plural de los sustantivos, el tiempo pasado de los verbos, el participio presente, o la formación de un sustantivo abstracto]. De este modo, la palabra *educat* pasa a ser la raíz que identifica también a *education*, *educating*, etc.; se considera que estos sufijos son formas gramaticales innecesarias, que no hacen más que complicar la elaboración de un diccionario.

En el General Inquirer, el diccionario identifica una raíz de palabra equiparándola unívocamente con una de sus entradas, y la asigna a una o más categorías denominadas "tipificadores" o "señaladores" [*tags*], que a partir de entonces se usan o bien juntamente con la palabra original o como representación singular. Varios niveles de señalamiento son posibles. El diccionario se desarrolla de forma independiente del texto y permanece fijo durante el análisis, aunque el usuario puede aumentarlo tras inspeccionar una lista de palabras remanentes que figuran en el texto pero no en el diccionario. Una peculiaridad del General Inquirer es que los señaladores o tipificadores se añaden simplemente a la cinta que contiene el texto original, de modo que puedan recuperarse las palabras originales. Una vez efectuado el señalamiento,

es posible llevar a cabo varias formas de análisis, incluidos el análisis de contingencia, el de aseveración evaluativa, diversos cómputos de frecuencias, y pueden obtenerse diversas listas, según los señaladores, las raíces de palabras o las oraciones que satisfacen determinados criterios, para su examen o para un análisis posterior.

La validez del General Inquirer procede de su diccionario, que representa una clasificación de vehículos-signos acorde con determinado criterio semántico. El diccionario primitivo recibió la fuerte influencia de la obra de Bales, *Interaction Process Analysis*, y llegó a ser conocido como el "Diccionario psicosociológico de Harvard". En su versión de 1965, tenía 3500 entradas y 83 señaladores. El hecho de que las palabras puedan interpretarse desde numerosas perspectivas (por ejemplo, "nitroglicerina" puede ser un explosivo, una sustancia química o una droga) dio origen a otros numerosos diccionarios, destinados, pongamos por caso, al análisis de los datos de una investigación por encuestas, a los estudios sobre el alcoholismo, al análisis del valor, aplicables a datos sobre necesidades de realización, al humor, a la mitología, a los titulares de periódicos y muchos otros. Los diccionarios tienden a ser específicos tanto de un discurso determinado como del ámbito de las inferencias pretendidas. Algunos se basan, sin lugar a dudas, en concepciones teóricas, mientras que otros se establecen después de una considerable labor experimental sobre la forma de agrupar las palabras por temas.

Aunque la concepción original del General Inquirer apuntaba al señalamiento de *todas* o casi todas las palabras del contenido de un texto (de ahí la necesidad de inspeccionar una lista de palabras remanentes), varios trabajos recientes sólo tienen en cuenta un conjunto menor de palabras, de significación teórica especial. Por ejemplo, ERTEL (1976) creó una escala de dogmatismo basándose en las frecuencias de dos clases fundamentales de palabras, "D+" y "D–". Los diccionarios tampoco necesitan atenerse a concepciones estrictamente semánticas. Así, Holsti (en STONE y otros, 1966) creó un diccionario que señala las palabras de acuerdo con las tres dimensiones del significado afectivo de Osgood: evaluación, intensidad y actividad. Su representación numérica se adecua a una teoría de las actitudes que, a su vez, justifica las inferencias realizadas acerca de las actitudes de un autor.

El tercero de los métodos de palabras separadas de su contexto queda ilustrado por un sistema que creó IKER (1975) y que deno-

minó WORDS. Fue desarrollado para analizar el contenido de datos procedentes de entrevistas psicoterapéuticas, tomando principalmente como modelo las experiencias previas con la preparación automática de resúmenes. Este sistema incluye muchos de los procedimientos del General Inquirer (por ejemplo, reducción de las palabras a sus raíces, diccionarios, listas de frecuencias), que no es preciso que reiteremos aquí. Un rasgo destacable es que evitó expresamente clasificar *a priori* las palabras, como hacen el método del diccionario y el del tesauro en el VIA y en el General Inquirer, respectivamente. En muchas circunstancias el investigador puede que no tenga ningún preconcepto acerca de cuáles son los términos teóricamente significativos, o tal vez no quiera establecer categorías sin examinar antes el texto que debe analizar. El sistema WORDS le ofrece un procedimiento por el cual ese esquema emerge del propio texto, sin incorporar la tendencia teórica que el investigador puede introducir en el análisis. Con los mismos fines ha sido utilizado el agrupamiento en conglomerados (DZIURZYNSKI, 1977).

Aunque la validez del sistema WORDS es discutible, pone de manifiesto la posibilidad de definir procedimientos de computación que permitan la aparición de categorías a partir de las propiedades mismas del texto. Estos procedimientos constituyen un paso adelante muy significativo hacia la generalización de los métodos del análisis de contenido, más allá de concepciones teóricas y particulares discursos.

Métodos de la inteligencia artificial

La labor desarrollada en el campo de la inteligencia artificial intenta principalmente representar la inteligencia humana mediante un ordenador. Tiene depositadas sus esperanzas en la comprensión de los procesos cognitivos humanos, el diseño de procedimientos mecánicos que ahorren al hombre procesos de control prescindibles o decisiones indeseables o difíciles, y la generalización de la inteligencia más allá de sus manifestaciones humanas. La inteligencia artificial comparte con el análisis de contenido por ordenador el interés por procesos como los de la comprensión del lenguaje, la formulación de inferencias a partir de formas simbólicas no lingüísticas de comunicación, y la adopción de decisiones inteligentes cuando se dispone de información incompleta.

Cuando los especialistas en inteligencia artificial prestaron

atención a los problemas de la traducción automática, varios de ellos entendieron que el análisis de contenido era una manera de traducir las expresiones del lenguaje natural en términos de un lenguaje científico, y en realidad la codificación efectuada en el análisis de contenido abarca procesos de esa índole. También la obtención de resúmenes de textos por ordenador se ajustaba a las concepciones tradicionales del análisis de contenido (por ejemplo, IKER, 1975). Cuando más tarde los especialistas en inteligencia artificial pasaron a ocuparse del diseño de sistemas de preguntas y respuestas, muchos estudiosos (por ejemplo, HAYS, 1969) vieron que también esto estaba implicado en el análisis de contenido. De hecho, todas las respuestas a las preguntas que formula un analista de contenido proceden del análisis selectivo de un texto. A diferencia de lo que ocurre en la recuperación de información, aquí la búsqueda de respuestas no se reduce al hecho de que un ítem cualquiera exista o no en una base de datos: requiere cierto nivel de comprensión. Además, las máquinas que procesan y responden instrucciones en lenguaje natural pueden considerarse analógicas con respecto a la labor de un analista de contenido, de quien se pretende que interprete sus datos y recomiende las medidas apropiadas (quizá sin el aspecto de "ejecución de órdenes" que es inherente a la comunicación entre el hombre y la máquina).

Más allá de estas analogías obvias, los enfoques de la inteligencia artificial aplicados al análisis de contenido por ordenador han sido más bien exploratorios, en lugar de constituir la fuente de sistemas analíticos a gran escala. Aunque quizá todavía sea demasiado prematuro para generalizar, en su mayoría estos enfoques incluyen un componente de análisis sintáctico seguido de un componente de análisis semántico, pero destacan que los resultados de sus componentes se homologan en un componente de la *lógica del discurso*, el cual representa la estructura lógica del mundo que se examina y puede almacenar y registrar conocimientos acerca de ese mundo.

Mientras que los métodos del diccionario y el tesauro pretendían obtener un amplio alcance reduciendo la complejidad de los fenómenos simbólicos a formas de discursos y a "mundos" bastante primitivos (en los cuales un determinado conjunto de unidades significa o se refiere a un conjunto menor, no ordenado, de cualidades o atributos), los empleados por la inteligencia artificial han tenido que hacer otras clases de concesiones al enfrentarse con complicaciones aún mayores. Por ejemplo, en una demostra-

ción de lo que implica "comprender el lenguaje natural", LINDSAY (1963) se dio por satisfecho con enunciados nucleares simples en "inglés básico"; no obstante, estos enunciados, almacenados en una memoria que representaba implícitamente la lógica del sistema de parentesco, conducen al mundo, bastante complejo, de las estructuras familiares actuales. En otro sistema, el "mundo" se limitaba a tablas de puntuaciones de los partidos de béisbol, pero el sistema aceptaba una amplia gama de preguntas en lenguaje natural, que podían responderse a partir de esas tablas. El ordenador analizaba la sintaxis y la semántica de esas preguntas (del tipo de "¿Cuántas veces el equipo de béisbol de los Red Sox batió al de los Yankees en 1958"?) y lograba encontrar respuestas a partir de los datos (GREEN y otros, 1963). Otra demostración simple, que trasciende los métodos del diccionario y del tesauro, es la que proporcionan KOPPELAAR y otros (1978). Estos autores programaron el "mundo" del analista como un modelo de simulación. Sus datos consistían en las dependencias causales percibidas entre las condiciones de trabajo en distintas organizaciones sociales, y fueron extraídos de transcripciones de entrevistas. La simulación exploraba todas las implicaciones funcionales de esos datos. HAYS (1969) sugirió que el "mundo" respecto del cual se analiza el contenido de un discurso debe estar organizado como una enciclopedia, y contener los presupuestos de las aseveraciones analizadas, los conocimientos previos sobre el tema y la estructura lógica de la situación.

Un avance interesante, y potencialmente importante, en los enfoques del análisis de contenido a través de la inteligencia artificial ha sido la creación de una serie de programas de ordenador capaces de representar computacionalmente discursos tales como descripciones de accidentes, interacciones sociales y conducta de los clientes de un restaurante, simulando de qué manera los individuos confieren sentido a estas elucidaciones verbales (que habitualmente dejan abandonados muchos detalles a la imaginación del lector). Este tipo de trabajos fue fruto del temprano interés por los modelos de cambio actitudinal (ABELSON, 1963) y aparecen en la obra de SCHANK y ABELSON (1977). Responde, en particular, a las críticas elevadas contra la lingüística computacional (que limitó en gran medida su atención a la comprensión de oraciones sueltas), tratando de representar todo un discurso en función de guiones, planes, temas y metas. Aunque la finalidad de los autores es fundamentalmente psicológica (a saber, comprender de qué mane-

ra los individuos almacenan y procesan el conocimiento que, en última instancia, puede llevarlos a controlar su conducta), las inferencias que han podido extraer de un texto son las mismas que varios analistas de contenido confiaron en obtener mediante técnicas no computacionales.

Teniendo en cuenta el estado actual de nuestros conocimientos, el análisis de contenido por ordenador probablemente alcance más éxito en los siguientes casos:

a) Cuando la atención se limita a los datos de un discurso particular, restringido en su vocabulario, sintaxis y semántica; el relativo éxito de los diccionarios especializados atestigua la sensatez de esta restricción.
b) Cuando la lógica del contexto de este discurso (causalidad, guiones, estructuras sociales) es, o bien enunciable explícitamente, o bien está implícita en la forma en que se representan los datos por vía computacional. Parece difícil computar significados e implicaciones y formular inferencias para todas las situaciones o mundos posibles: sólo aquellos que puedan formalizarse en un grado razonable pueden proporcionar resultados válidos.
c) Cuando los conocimientos y presupuestos *a priori* relevantes para el discurso pueden incorporarse al análisis, ya sea en forma de datos adicionales, o en forma de la lógica del contexto; en una y otra forma, deben contribuir a las inferencias deseadas.

12. Fiabilidad

En este capítulo se efectúa una amplia reseña de los procedimientos mediante los cuales puede y debe evaluarse la fiabilidad en el análisis de contenido. Se exponen diseños para el examen de la estabilidad, la reproducibilidad y la exactitud. También se desarrollan y se debaten las técnicas computacionales, los patrones comparativos y los procedimientos de diagnóstico.

Si se pretende que los resultados de una investigación sean válidos, tanto los datos sobre los que se basan, como los individuos que participan en su análisis y los procesos mediante los cuales se llega a dichos resultados, deben ser fiables. La fiabilidad es una condición necesaria —aunque no suficiente— de la validez.

La evaluación de la fiabilidad actúa como una importante salvaguardia contra la contaminación de los datos científicos por efectos ajenos a las finalidades de la observación, medición y análisis. Tal como sostienen KAPLAN y GOLDSEN:

> La importancia de la fiabilidad procede de la seguridad que ofrece en cuanto a que los datos han sido obtenidos con independencia del suceso, instrumento o persona que los mide. Por definición, los datos fiables son aquellos que permanecen constantes en todas las variaciones del proceso de medición (1965, págs. 83-84).

La fiabilidad mide el grado en el cual cualquier diseño de investigación, o cualquiera de sus partes, o cualquiera de los datos resultantes, representan variaciones en los fenómenos reales, en lugar de representar las circunstancias extrínsecas de la medición, las idiosincrasias ocultas de cada uno de los analistas o las tendencias subrepticias de un procedimiento.

Para verificar la *fiabilidad* es esencial que haya cierta duplicación de los esfuerzos. Un procedimiento fiable es aquel que rinde los mismos resultados para los mismos conjuntos de fenómenos, independientemente de las circunstancias de su aplicación. Para verificar la *validez*, en cambio, los resultados de un procedimiento deben ajustarse a lo que, según se sabe, es "verdadero" o lo que ya se presume válido. De esto se desprende que, si por un lado la fiabilidad asegura que los resultados analíticos representan algo real, la validez asegura que dichos resultados representan lo que pretenden representar. En el análisis de contenido, la fiabilidad y la validez están relacionadas entre sí por las dos proposiciones siguientes:

La fiabilidad establece límites a la validez potencial de los resultados de la investigación. Sin lugar a dudas, ante un proceso que carezca por completo de fiabilidad, cuyos resultados son simples sucesos azarosos, no puede esperarse que las conclusiones sean válidas en un grado mayor que el azar. Aunque en la práctica rara vez se está ante procesos de fiabilidad nula, la probabilidad de un resultado válido no puede superar la probabilidad con la cual puede obtenerse ese resultado en el proceso de los análisis repetidos.

La fiabilidad no garantiza la validez de los resultados de la investigación. Dos jueces con iguales prejuicios pueden estar de acuerdo en lo que ven, pero estar equivocados en relación con todos los demás criterios. Un procedimiento analítico puede ser determinista, como un programa de ordenador, y sin embargo no tener nada en común con el contexto del cual proceden los datos, y reiterar sus errores una y otra vez. De modo, pues, que una alta fiabilidad no proporciona seguridad alguna de que los resultados sean en realidad válidos.

Podría añadirse a las dos anteriores una tercera proposición, que no tiene, sin embargo, su fuerza definitoria: *La fiabilidad a menudo entorpece la validez*. Este hecho empírico procede en par-

te de las dificultades efectivas que plantea el análisis sistemático de formas simbólicas complejas, y en parte de la tendencia natural del investigador a mejorar la calidad de los datos que puede medir con más facilidad. Un ejemplo, examinado por HOLSTI (1969), es el estudio de la conciencia nacional de las trece colonias norteamericanas llevado a cabo por MERRITT (1966) sobre la base de los relatos periodísticos. Dado que es difícil determinar las unidades, registrar y comparar los temas en que aparecen manifestados esos sentimientos, temas que por lo tanto probablemente resulten poco fiables, el análisis enumeró en vez de ello los nombres de lugares. De hecho, el cómputo de palabras rara vez plantea problemas de fiabilidad, aunque su validez sea incierta.

En lo que sigue vamos a diferenciar las condiciones aptas para generar datos adecuados para las pruebas de fiabilidad, definiremos una medida general del acuerdo para los estudios de fiabilidad, discutiremos los criterios para aceptar los datos como suficientemente fiables con vistas a un análisis posterior, y ofreceremos varios procedimientos diagnósticos que son útiles para diversos tipos de control de calidad, particularmente durante la fase de verificación previa del análisis de contenido.

DISEÑOS PARA VERIFICAR LA FIABILIDAD

En realidad, con el término "fiabilidad" se designan como mínimo tres tipos distintos de fiabilidad, y conviene aclarar cuál de ellos se aplica en una situación determinada. Por ejemplo, un codificador puede replicar lo que ya ha hecho antes y, al comprobar que no existen desviaciones fundamentales entre uno y otro procedimiento, llegar a la conclusión de que sus datos son fiables. Compárese esta situación con aquella otra en que la asignación de unidades a categorías por un codificador se coteja con la que efectúa otro codificador. También en este caso la ausencia de variabilidad es un indicador de la fiabilidad de los datos, pero la prueba es mucho más eficaz, porque es sensible a algo más que el "ruido" interno o las incongruencias de un solo codificador. La diferencia entre las dos situaciones es de diseño, es decir, depende de la manera en que se obtienen los datos sobre la fiabilidad, y no de la manera en que se evalúan los dos conjuntos de datos.

Distinguiremos, pues, estos tres aspectos:

- estabilidad
- reproducibilidad
- exactitud

La *estabilidad* de un proceso es el grado en el que permanece invariante o sin modificaciones a lo largo del tiempo. La estabilidad se pone de manifiesto en las condiciones de un test-retest, como por ejemplo cuando a un mismo codificador se le solicita codificar dos veces un conjunto de datos, en momentos distintos. Las discrepancias entre las dos maneras que se describen, codifican y miden las unidades, o en que se asignan a valores numéricos o categorías, reflejan las inconsistencias o el "ruido" propio de cada observador, los cambios cognitivos que tuvieron lugar *dentro* de él o su dificultad para interpretar las instrucciones de registro. La estabilidad es la forma más débil de fiabilidad, y no debe pensarse en ella como indicador exclusivo de la aceptabilidad de los datos de un análisis de contenido para la posterior formulación de inferencias y análisis. A la estabilidad se la conoce también como "fiabilidad del observador" o, simplemente, como "congruencia" o "consistencia".

La *reproducibilidad* de un proceso es el grado en que puede recrearse en circunstancias diferentes, en otros lugares y con la intervención de codificadores distintos. Para establecer la reproducibilidad, los datos deben obtenerse en condiciones de test-test. Un ejemplo sería cuando dos o más individuos aplican las mismas instrucciones de registro de forma independiente al mismo conjunto de datos. Las discrepancias en cuanto al modo en que estos individuos registran los datos reflejan tanto las incongruencias de cada uno como las diferencias que hay entre ellos en su manera de interpretar las instrucciones. A la reproducibilidad se la denomina también "fiabilidad entre los codificadores", "acuerdo intersubjetivo" o "consenso" logrado por los observadores.

La *exactitud* de un proceso es el grado en el que se ajusta funcionalmente a un criterio o patrón conocido, o el grado en que sus resultados son lo que se planeaba que fueran. Para establecer la exactitud, los datos se obtienen en condiciones de test-norma; por ejemplo, cuando el trabajo de un codificador o un instrumento de

medida se compara con lo que, según se sabe, es el trabajo o medida correctos. En tal caso, las diferencias entre las dos formas de registro reflejan las inconsistencias de cada observador, las discrepancias entre ellos, así como las desviaciones sistemáticas respecto de la norma. La exactitud es, de todas las pruebas de fiabilidad de que se dispone, la más eficaz, sólo superada por una medida de la validez, que se ajusta a este mismo diseño, salvo que además debe saberse que la norma es verdadera, lo cual en ningún momento puede decirse de las evaluaciones de la fiabilidad.

En la tabla 2 se ilustran las diferencias entre estos tres tipos de diseños para verificar la fiabilidad.

Tabla 2
Tipos de fiabilidad

Tipos de fiabilidad	Diseños para verificar la fiabilidad	Errores evaluados	Intensidades relativas
Estabilidad	test-retest	incongruencias del observador	el menos eficaz
Reproducibilidad	test-test	incongruencias del observador y desacuerdos entre los observadores	
Exactitud	test-norma	incongruencias del observador, desacuerdos entre los observadores y desviaciones sistemáticas respecto de una norma	el más eficaz

Las mediciones de la exactitud han demostrado ser útiles para verificar el trabajo de los codificadores en el curso de su capacitación, cuando se ha establecido previamente una norma o un grupo de expertos la ha determinado. También resulta claro el significado de la exactitud cuando los datos se manejan de una forma manual rutinaria o por ordenador; pero en la mayoría de las situaciones, cuando las observaciones, contenidos de los mensajes y textos son codificados según las categorías de un lenguaje de datos, rara vez se dispone de normas con respecto a las cuales establecer dicha exactitud. Por consiguiente, en el análisis de contenido es poco realista insistir en este criterio de fiabilidad, el más potente de todos. *Los datos deben ser por lo menos reproducibles* por investigadores

independientes en distintos lugares y momentos, utilizando las mismas instrucciones para codificar las mismas series de datos.

Adviértase que la reproducibilidad exige que los codificadores sean *independientes*. Lamentablemente, la bibliografía sobre análisis de contenido informa de numerosas violaciones de este requisito. Por ejemplo, puede solicitarse a varios codificadores que debatan entre sí cada unidad de observación y lleguen a un acuerdo respecto de cuál sería la mejor categoría descriptiva. No obstante, esta práctica no asegura la reproducibilidad ni tampoco pone de relieve su grado. Si los procesos grupales pueden suprimir las idiosincrasias personales, la comunicación entre los codificadores introduce otros errores; y como está previsto que el proceso dé por resultado una sola puntuación por unidad (todos los individuos participantes funcionan como un observador único), los datos no suministran indicio alguno en cuanto a la fiabilidad. En otro ejemplo, aunque los codificadores trabajan de modo independiente, se les permite comunicarse entre sí cuando surgen problemas; la comunicación influye invariablemente en la codificación, consiguiendo que se alcance un mayor grado de acuerdo, y *es probable que esta falta de independencia haga que los datos parezcan más fiables de lo que son.*

Otra costumbre, presuntamente destinada a eludir los problemas de la falta de fiabilidad, es tomar los promedios o los juicios mayoritarios como valores verdaderos cuando surgen discrepancias entre observadores independientes. Aunque es cierto que estos datos contienen evidencias acerca de la reproducibilidad *antes* de eliminar la variabilidad "indeseable", la discrepancia nada nos dice acerca de quién está en lo cierto y quién está equivocado; por lo tanto, ni la media ni la moda tienen el poder de mejorar la fiabilidad de los datos por medios computacionales.

Una práctica más engañosa todavía consiste en admitir únicamente en la investigación aquellos datos respecto de los cuales los codificadores independientes han alcanzado un acuerdo perfecto. Aquí la tendencia es doble: el procedimiento no impide que ciertos acuerdos azarosos se introduzcan en los datos (pueden representar incluso el 50% del total de las unidades) y desvía los datos hacia aquello que es fácilmente codificable. En toda investigación científica los datos deben escogerse por ser representativos del fenómeno que interesa conocer, y no por adecuarse a las necesidades de un método determinado.

Un diseño que se utiliza erróneamente para evaluar la fiabilidad

en el análisis de contenido es la técnica de la división en mitades. En este caso una muestra de unidades de registro se divide en dos partes iguales; si las diferencias entre ambas no son estadísticamente significativas se considera que los datos son fiables; de lo contrario, se considera que no lo son. Sin embargo, con esta técnica sólo se pone a prueba la homogeneidad estadística dentro de una muestra. Puede establecerse así si dicha muestra es lo bastante amplia como para representar las características de la población de la cual ha sido extraída (véase el capítulo 6, "Muestreo"); pero no se verifica con ello si las unidades, tal como han sido registradas, son portadoras de información individualmente, si han sido descritas de forma consistente y si, en consecuencia, pueden considerarse descriptivas de algún fenómeno real.

Acuerdo

Los *datos sobre fiabilidad* exigen que dos codificadores, como mínimo, describan de forma independiente un conjunto posiblemente amplio de unidades de registro en los términos de un lenguaje de datos común. Estos términos pueden ser categorías descriptivas, valores numéricos de una variable o complejos esquemas clasificatorios.

La fiabilidad se expresa como una función del acuerdo alcanzado entre los codificadores sobre la asignación de las unidades a las diversas categorías. Si dicho acuerdo es total para todas las unidades, está garantizada la fiabilidad; si no es mayor que el correspondiente al azar (lo cual podría suceder si los codificadores no intentan examinar las unidades, o si en lugar de ello lanzan un dado para resolver las asignaciones a categorías), la fiabilidad es nula. Ya sea que la fiabilidad adopte el aspecto de la estabilidad, la reproducibilidad o la exactitud, siempre se reduce a medir el acuerdo alcanzado por los observadores, codificadores o jueces al procesar de forma independiente la información científica.

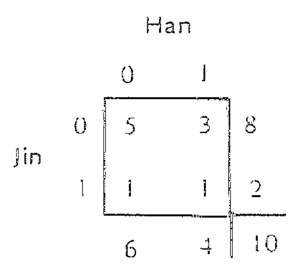
Ejemplo

Veamos un ejemplo muy simple. Supongamos que alguien quiere estudiar la percepción que existe de la política exterior norteamericana en China tal como se refleja en los órganos de prensa de este país, e instruye a dos hablantes del idioma chino para que identifiquen la

presencia o ausencia de las referencias a dichas percepciones marcando un documento con los dígitos "1" y "0", respectivamente. Los datos de esta situación test-test podrían representarse así:

documento	1	2	3	4	5	6	7	8	9	10
Jin	0	1	0	0	0	0	0	0	1	0
Han	0	1	1	0	0	1	0	1	0	0

Los datos de este cuadro muestran que los dos sujetos, que aquí llamamos Jin y Han, coinciden en seis de un total de diez asignaciones de noticias a las categorías (o sea, en el 60%). Ahora bien, un porcentaje así no nos dice mucho. No nos dice nada acerca de si esta cifra es grande o pequeña, o de qué manera puede compararse con lo que sucedería por obra del azar. Para tener en cuenta estos interrogantes, colocamos los datos de la matriz 2 x 10 en una tabla de contingencia de frecuencias de 2 x 2.



		Han 0	Han 1	
Jin	0	5	3	8
	1	1	1	2
		6	4	10

Para obtener una tabla semejante, en la que el acuerdo es producto del azar, debe contarse con algún conocimiento acerca de la manera en que las referencias a la política exterior norteamericana se distribuyen en los órganos de prensa chinos. Aunque ignoramos esta distribución, podemos estimarla a partir de la experiencia de ambos codificadores. Como Jin sostiene que ha identificado dos de un total de diez noticias, y Han sostiene que identificó cuatro de un total de diez, ambos en conjunto sostienen que identificaron seis de un total de veinte (o sea, el 30%) del total de noticias, con referencias a la política exterior norteamericana. Esta es exactamente la proporción de "1" en la matriz de 2 x 10 de datos sobre fiabilidad. Supóngase ahora que los catorce "0" y los seis "1" se colocan dentro de una caja de la cual dos individuos extraen papeles al azar. Al extraer dos "1", el primer sujeto habrá extraído un "1", en seis de un total de veinte casos; y luego de

extraer ese "1", y por lo tanto sacarlo de la caja, el segundo codificador extraerá un "1" equiparable en cinco de un total de 19 casos. La probabilidad de que concuerden entre sí acerca de la presencia de una referencia a la política exterior norteamericana por mero azar es, pues, el producto de estas dos probabilidades: (6/20) (5/19) = 0,08. Multiplicada por la cantidad de unidades de registro (diez en este caso), la frecuencia prevista es 0,8 o menor que la unidad. Las otras frecuencias previstas se obtienen mediante un razonamiento semejante.

Una tabla en que el acuerdo es perfecto sólo tiene valores en la diagonal principal, o sea, la que va del casillero 0-0 al casillero 1-1, y todo lo que queda fuera de ella es nulo. En una tabla de este tipo, los valores corresponden a las proporciones estimadas de la población. Así pues, las tres tablas a las que debe remitir una medida adecuada de acuerdo son:

7	
	3

acuerdo máximo

5	3
1	1

co-ocurrencias observadas

4,8	2,2
2,2	0,8

acuerdo por azar

La medida más convincente para expresar la cantidad de acuerdo en los datos sobre fiabilidad es el grado de semejanza entre la tabla de co-ocurrencias observadas y la tabla de acuerdo máximo, más que el grado en que el acuerdo es meramente producto del azar. Esto se expresa en la siguiente fórmula, donde el signo α es el coeficiente de acuerdo buscado:

$$\begin{matrix}\text{co-ocurrencias}\\ \text{observadas}\end{matrix} = \alpha \begin{pmatrix}\text{acuerdo}\\ \text{máximo}\end{pmatrix} + (1-\alpha)\begin{pmatrix}\text{acuerdo}\\ \text{por azar}\end{pmatrix}$$

Si la tabla de co-ocurrencias observadas es igual a la tabla de acuerdo máximo, entonces $\alpha = 1$, y si es igual a la tabla de acuerdo por azar, entonces $\alpha = 0$. Tras algunas operaciones algebraicas se llega a la siguiente fórmula para la expresión del coeficiente:

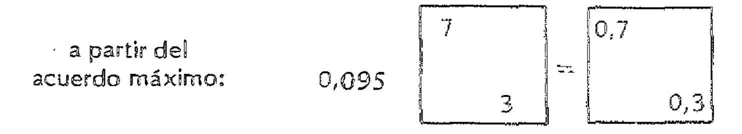
$$\alpha = 1 - \frac{\text{discrepancia observada}}{\text{discrepancia prevista}}$$

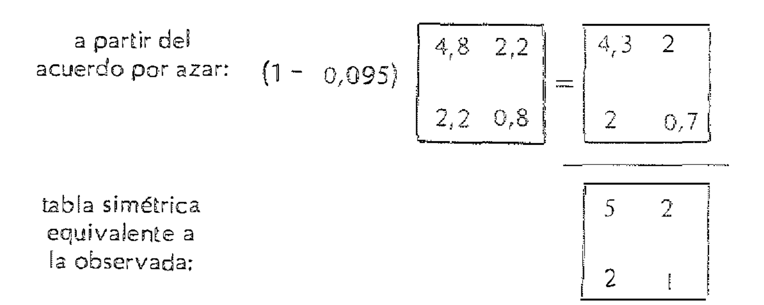
En la tabla de 2 x 2, las discrepancias aparecen en los casilleros 0-1 y 1-0, la discrepancia observada es 3 + 1 = 4 y la discrepancia prevista es 2,2 + 2,2 = 4,4, de modo que el acuerdo de los datos sobre fiabilidad es:

$$\alpha = 1 - \frac{(3 + 1)}{(2,2 + 2,2)} = 0,095$$

Vemos que el acuerdo resulta estar apenas un diez por ciento por encima del obtenido al azar. Podría decirse que la actuación en estos dos codificadores equivale a haber leído e identificado correctamente la referencia a la política exterior norteamericana sólo en un diez por ciento de todas las noticias, y luego haber lanzado los dados para determinar las categorías de las restantes. Tal vez uno o ambos codificadores no conocían suficientemente la lengua china, o tal vez no lograron entender las instrucciones que se proporcionaron en inglés. Quizás estas instrucciones eran ambiguas o no resultaban apropiadas para los datos de que se disponía. Sea como fuere, estos datos deben rechazarse por no ser reproducibles. En estas circunstancias, las conclusiones a las que ellos darían lugar serían en gran parte engañosas o verdaderas sólo por azar. El hecho de que el 60% del total de asignaciones a categorías coincidiesen, no tiene ningún tipo de significado. Esto debe tomarse como advertencia contra el uso de un acuerdo porcentual como criterio de fiabilidad. Aunque se menciona a menudo en la bibliografía sobre el análisis del contenido, es totalmente engañoso.

Según lo anterior, el coeficiente de acuerdo indica de qué manera es posible reconstruir la tabla original (u otra equivalente en cuanto a la distribución de los acuerdos) a partir de las dos tablas de acuerdo máximo y de acuerdo por azar, respectivamente:





donde vale la pena advertir que las tablas siguientes son todas equivalentes entre sí en cuanto al grado de acuerdo/discrepancia que incluyen:

5	4
0	1

5	3
1	1

5	2
2	1

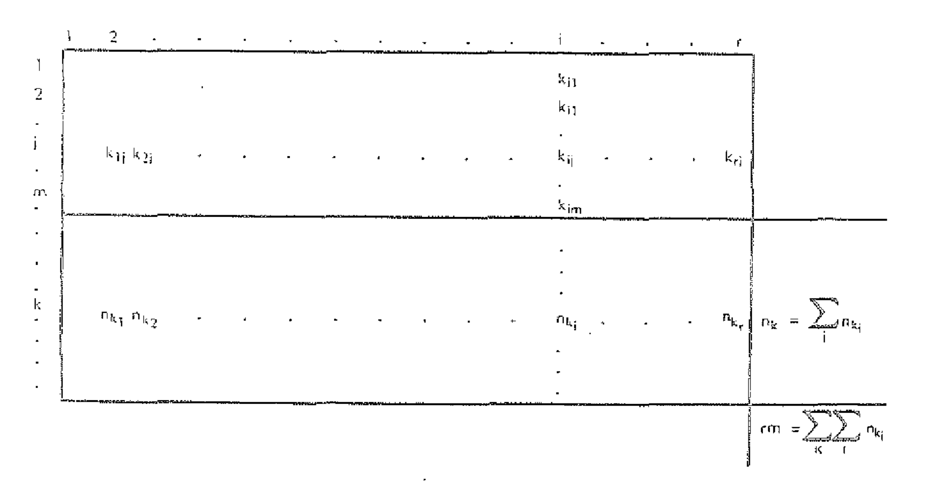
5	1
3	1

5	0
4	1

El ejemplo que hemos dado es demasiado simple, y el coeficiente de acuerdo así establecido, muy limitado. En primer lugar, porque los datos sólo abarcaban dos codificadores, mientras que a menudo es preciso evaluar los acuerdos entre muchos observadores; y en segundo lugar, porque sólo se ocupaban de elecciones dicotómicas, mientras que en el análisis de contenido los datos pueden incluir muchísimas distinciones y métricas o escalas complejas. Teniendo en cuenta este sencillo ejemplo, pasaremos ahora a enunciar en términos más generales el coeficiente de acuerdo.

Forma canónica de los datos sobre fiabilidad

En lugar de dos codificadores, veamos ahora un caso en que se trate de m, y en vez de una distinción dicotómica, consideremos un sistema de categorías o valores de los cuales k_{ij} sea el asignado a la unidad de registro i-ésima por el juez, observador o codificar j-ésimo. Los datos sobre fiabilidad adoptan entonces la forma de una matriz m x r, a la que llamaremos su *forma canónica*:



En la parte inferior de esta matriz se describen las frecuencias n_{ki} con las cuales la categoría k aparece en la columna i, y las frecuencias n_k de la categoría k en el cuadro en su conjunto.

Diferencias y discrepancias

Para evaluar las discrepancias que puedan presentarse en una tabla de esta índole, es importante conocer el grado de diferencia entre dos categorías o valores, lo cual depende de la naturaleza de la métrica en cuestión. La idea es relativamente simple: cuando las categorías de una variable no están ordenadas (como sucede con los nombres de personas, lugares de nacimiento u ocupaciones), no importa cuáles sean las dos categorías que puedan haberse confundido: toda falta de coincidencia es tan buena o tan mala como cualquier otra. En cambio, si las categorías están ordenadas (como sería el caso de las medidas de tiempo, longitud o de intensidad actitudinal), la distancia que separa a los valores no coincidentes puede generar una diferencia considerable. Sin duda, la diferencia entre uno y veinte es mayor que la diferencia entre uno y dos si nos estamos ocupando de kilómetros, unidades monetarias o edades. Como ya examinamos al tratar los lenguajes de datos, existen muchas clases de métricas. La edad, por ejemplo, constituye una escala de razones. Así, conjeturar con precisión de un año la edad de una persona de alrededor de ochenta años es un

buen acierto, en tanto que no lo sería conjeturar con esa misma precisión la edad de una criatura de pocos años. En las escalas de razones, las distancias dependen de la separación que exista entre ellas y el punto cero. Las definiciones siguientes tienen en cuenta las escalas nominales, ordinales, de intervalos y de razones.

Si las dos categorías o valores de una variable son b y c, las diferencias entre ellas se plantearían de la siguiente manera:

Para las *escalas nominales* o variables no ordenadas,

$$d_{bc} = \begin{cases} 0 & \text{si } b = c \\ 1 & \text{si } b \neq c \end{cases}$$

Para las *escalas ordinales*,

$$d_{bc} = \left(\sum_{k>b} \frac{n_k}{rm} \quad \sum_{k<b} \frac{n_k}{rm} + \sum_{k<c} \frac{n_k}{rm} - \sum_{k>c} \frac{n_k}{rm} \right)^2$$

Para las *escalas de intervalos*,

$$d_{bc} = (b-c)^2$$

Para las *escalas de razones*,

$$d_{bc} = \left(\frac{b-c}{b+c}\right)^2$$

Hay otras diferencias posibles, de las que no nos ocuparemos aquí. Ahora bien, en la matriz m X r de datos sobre fiabilidad en su forma canónica, cada columna contiene m entradas, que son las categorías o valores asociados con una unidad. Dentro de cada columna es posible formar m (m-1) pares sin prestar atención al orden. Si n_{b_i} es la cantidad de categorías b en la columna i, y n_{c_i} es la cantidad de categorías c en la columna i, *el número de b-c pares en la columna i* es $n_{b_i} n_{c_i}$ si $b \neq c$, y $n_{b_i} (n_{c_i} - 1)$ si $b = c$. Por supuesto, la suma da

$$m(m-1) = \sum_{b} \sum_{c} n_{b_i} (n_{c_i} - \Delta_{bc}) \text{ donde } \Delta_{bc} = \begin{cases} 0 & \text{si } b \neq c \\ 1 & \text{si } b = c \end{cases}$$

donde la delta de Kronecker asegura que las entradas no se comparen entre sí. Como existen r columnas, la *discrepancia observada* puede expresarse entonces como la diferencia media dentro de la columna (unidad):

$$D_o = \frac{1}{r}\sum_i \sum_b \sum_c \frac{n_{b_i} n_{c_i}}{m(m-1)} d_{bc}.$$

Prescindiendo de la referencia a las columnas, *el número de posibles parejas b-c en la matriz total es* $n_b n_c$, que suma

$$rm(rm-1) = \sum_b \sum_c \left(\sum_i n_{b_i} \left(\sum_i n_{c_i} - \Delta_{bc} \right) \right) = \sum_b \sum_c n_b (n_c - \Delta_{bc}).$$

Y la *discrepancia prevista* vuelve a expresarse como una diferencia media, pero ahora de todos los pares posibles en la matriz total:

$$D_c = \sum_b \sum_c \frac{n_b n_c}{rm(rm-1)} d_{bc}.$$

Coeficiente de acuerdo para datos canónicos

En función de los datos sobre fiabilidad en su forma canónica, el *coeficiente de acuerdo* se define por

$$\alpha = 1 - \frac{D_o}{D_c}$$

$$= 1 - \frac{\frac{1}{r}\sum_i \sum_b \sum_c \frac{n_{b_i} n_{c_i}}{m(m-1)} d_{bc}}{\sum_b \sum_c \frac{n_b n_c}{rm(rm-1)} d_{bc}}$$

$$= 1 - \frac{rm-1}{m-1} \frac{\sum_i \sum_b \sum_{c>b} n_{b_i} n_{c_i} d_{bc}}{\sum_b \sum_{c>b} n_b n_c d_{bc}}$$

La primera expresión presenta la idea general, poniendo en evidencia que cuando no hay discrepancia en los datos, el numerador

es 0 y el coeficiente de acuerdo es igual a 1, lo cual indica que la fiabilidad es total; mientras que cuando la discrepancia observada es igual a lo que se espera por azar, el coeficiente de acuerdo es 0, lo cual indica que la fiabilidad es nula.

En la segunda expresión, las discrepancias observadas y previstas se introducen como diferencia media. La tercera expresión ofrece una fórmula de cálculo más conveniente.

Es importante advertir que cuando el número de codificadores es exactamente dos, las categorías de las variables no están ordenadas (escala nominal) y la muestra es de enorme tamaño, nuestro coeficiente de acuerdo es igual a la pi de SCOTT (1955):

$$\text{pi} = \frac{\text{\% de coincidencias observadas} - \text{\% de coincidencias previstas}}{100 - \text{\% de coincidencias previstas}}$$

$$= 1 - \frac{100 - \text{\% de coincidencias observadas}}{100 - \text{\% de coincidencias previstas}}$$

En estas fórmulas, y a diferencia de lo que ocurre con el coeficiente kappa de COHEN (1960), el porcentaje de coincidencias previstas se calcula por la proporción con que se utiliza una categoría, tomando ambos codificadores. La expresión (100 - % de coincidencias) es de discrepancia, y pone de manifiesto la conformidad entre la "pi" y nuestra fórmula. Nuestro coeficiente permite hacer correcciones para muestras de pequeño tamaño. Para dos codificadores y juicios en escalas de intervalos, nuestro coeficiente es asimismo igual al coeficiente de correlación intraclase. Para más de dos codificadores, formaliza un método sugerido por SPIEGELMAN y otros (1953a), que implica ordenamientos subjetivos por rangos de las pautas de discrepancia entre individuos que asignan las categorías por escala nominal. Así pues, nuestro coeficiente es una generalización para muchos codificadores, muchas clases de órdenes en los datos (métricas) y muchos tamaños de muestra. Suministra una medida uniforme que es comparable en numerosas situaciones.

Un ejemplo hipotético de datos sobre fiabilidad obtenidos para tres codificadores ilustrará el cálculo:

		1	2	3	4	5	6	7	8	9	
unidad i:		1	2	3	4	5	6	7	8	9	
codificador:	A	1	1	2	4	1	2	1	3	2	
	B	1	2	2	4	4	2	2	3	2	
	C	1	2	2	4	4	2	3	3	2	
n_{1i} =		3	1			1		1			n_1 = 6
n_{2i} =			2	3			3	1		3	n_2 = 12
n_{3i} =								1	3		n_3 = 4
n_{4i} =					3	2					n_4 = 5

Para unidades con acuerdo perfecto, las diferencias dentro de las columnas son cero en todos los casos. De este modo, las unidades 1, 3, 4, 6, 8 y 9 no contribuyen a la discrepancia observada. Presumiendo que los valores en esta variable no estén ordenados (escala nominal), la diferencia entre los pares no coincidentes es 1, y 0 en caso contrario. El valor del numerador del coeficiente en la fórmula de cálculo es

$$\sum_i \sum_b \sum_{c>b} n_{bi}\, n_{ci}\, d_{bc} = n_{12}\, n_{22} + n_{15}\, n_{45} + n_{17}\, n_{27} + n_{17}\, n_{37} + n_{27}\, n_{37}$$
$$= 1\cdot 2 + 1\cdot 2 + 1\cdot 1 + 1\cdot 1 + 1\cdot 1$$
$$= 7$$

y el valor del denominador es

$$\sum_b \sum_{c>b} n_b\, n_c\, d_{bc} = n_1\, n_2 + n_1\, n_3 + n_1\, n_4 + n_2\, n_3 + n_2\, n_4 + n_3\, n_4$$
$$= 6\cdot 12 + 6\cdot 4 + 6\cdot 5 + 12\cdot 4 + 12\cdot 5 + 4\cdot 5$$
$$= 254$$

si r =9 unidades y m = 3 codificadores, el coeficiente da

$$\alpha = 1 - \frac{9\cdot 3-1}{3-1}\ \frac{7}{254} = 0{,}642.$$

Este acuerdo indicaría que, en alrededor del 65% de los casos,

las co-ocurrencias observadas son explicables por la pauta de acuerdo perfecto más que por lo que sería previsible como consecuencia del azar; en suma, que las co-ocurrencias observadas se presentan en un 65% más de ocasiones que en el caso de regir el azar.

Si este caso resulta aún desconcertante, tómese el ejemplo del comienzo de este apartado, que implicaba las elecciones dicotómicas 0 ó 1 efectuadas por m =2 codificadores sobre r = 10 unidades. La diferencia entre las categorías no coincidentes vuelve a ser $d_{01} = d_{10} = 1$, y $d_{00} = d_{11} = 0$. En cada columna, o bien hay una falta de coincidencia o no la hay, de modo que la discrepancia observada expresa simplemente la cantidad de faltas de coincidencias como proporción de la cantidad total de pares en las columnas:

$$D_o = \frac{1}{r} \sum_i \sum_b \sum_c \frac{n_{bi}\, n_{ci}}{m(m-1)} d_{bc} = \frac{\text{número de unidades no coincidentes}}{r}$$

y en este caso la discrepancia prevista puede simplificarse análogamente de este modo:

$$D_e = \sum_b \sum_c \frac{n_b\, n_c}{rm(rm-1)} d_{bc} = \frac{n_0\, n_1}{r(2r-1)}$$

de manera que el coeficiente de acuerdo para un acuerdo bivariado (m =2 codificadores solamente) sobre elecciones dicotómicas (0 ó 1) se reduce a:

$$\alpha = 1 - \frac{D_o}{D_e} = 1 - \frac{(2r-1) \text{ número de unidades no coincidentes}}{n_0\, n_1}$$

Si las unidades no coincidentes fueran cuatro en este ejemplo, el valor del coeficiente sería:

$$\alpha = 1 - \frac{(2 \cdot 10 - 1)\, 4}{14 \cdot 6} = 0{,}095$$

que es igual al que habíamos obtenido anteriormente.

Datos sobre fiabilidad en forma de una matriz de coincidencia

Con frecuencia es más conveniente desde el punto de vista conceptual examinar los datos sobre fiabilidad no en su forma canónica, sino más bien en forma de una tabla simétrica de co-ocurrencias denominada *matriz de coincidencia.* No debe confundirse esta matriz con la matriz de contingencia, que no es simétrica y que sólo es aplicable a dos codificadores. En la matriz de coincidencia se ingresan los *números de pares*, x_{bc}, de los valores b y c de todas las unidades. Puede construirse a partir de la forma canónica mediante la fórmula siguiente:

$$x_{bc} = \sum_{i} n_{b_i} (n_{c_i} - \Delta_{bc}) \quad \text{donde nuevamente} \quad \Delta_{bc} = \begin{cases} 0 & \text{si } b \neq c \\ 1 & \text{si } b = c \end{cases}$$

La tabla adopta el siguiente aspecto, obteniendo los marginales mediante adición, según se indica:

$$\cdots\ x^{\cdot p}\ \cdots \qquad x^{\cdot\cdot} = \sum^{p} \sum^{c} x^{pc} = \omega(\omega - 1)$$

$$c \quad \cdots\ x^{pc}\ \cdots \qquad x^{\cdot c} = \sum^{p} x^{pc} = u^{c}(\omega - 1)$$

$$p$$

En esta matriz, los pares coincidentes ocupan las diagonales, mientras que los no coincidentes se distribuyen en torno a la diagonal. Como la diferencia entre los valores coincidentes es siempre igual a 0, la discrepancia observada es la diferencia media de los pares que no ocupan la diagonal. La discrepancia prevista calcula ese promedio bajo el supuesto de la independencia estadística ($p_{bc} = p_b\, p_c$).

Coeficiente de acuerdo para las formas matriciales

Teniendo en cuenta esta representación de los datos sobre fiabilidad mediante la matriz de coincidencia, el *coeficiente de acuerdo* se hace algo más sencillo de calcular:

$$\alpha = 1 - \frac{D_o}{D_e}$$

$$= 1 - \frac{\sum_b \sum_c \frac{x_{bc}}{x_{..}} d_{bc}}{\sum_b \sum_c \frac{x_{b.}\, x_{.c}}{x_{..}\,(x_{..} - m + 1)} d_{bc}}$$

$$= 1 - \frac{rm-1}{rm} \frac{\sum_b \sum_c \frac{x_{bc}}{x_{..}} d_{bc}}{\sum_b \sum_c \frac{x_{b.}\, x_{.c}}{x_{..}\, x_{..}} d_{bc}}$$

$$= 1 - (x_{..} - m + 1) \frac{\sum_b \sum_{c>b} x_{bc}\, d_{bc}}{\sum_b \sum_{c>b} x_{b.}\, x_{.c}\, d_{bc}}$$

La primera expresión es, nuevamente, la forma general en que se ingresa la discrepancia observada y la prevista a fin de obtener la segunda expresión. La tercera pone de relieve cómo se corrige el coeficiente de acuerdo cuando las muestras son pequeñas. Si el número de codificadores, m, o el de unidades, r, resultan muy grandes, el factor (rm - 1)/rm se aproxima a la unidad y puede dejarse de lado. Si las muestras fueran de tamaño reducido, su efecto sería reducir el coeficiente, que en este caso es corregido. Separando este factor de corrección, el resto, según puede apreciarse, expresa las diferencias observadas tal como se encuentran en la matriz, y las diferencias previstas como se presentarían en el supuesto de independencia estadística. La cuarta expresión proporciona una fórmula de cómputo conveniente.

En un ejemplo correspondiente al tercer codificador y a la novena unidad, la matriz de coincidencia de las co-ocurrencias observadas sería la siguiente:

categorías:

	1	2	3	4	
1	6	3	1	2	12
2	3	20	1		24
3	1	1	6		8
4	2			8	10
	12	24	8	10	54

Nótese que esta tabla es simétrica. No reconoce el hecho de que la contribución individual de los codificadores y sus márgenes se ven incrementados con un factor igual a (m - 1), o sea, el número de grados de libertad en cada columna (unidad).

El número observado de pares no coincidentes en uno de los triángulos que están fuera de la diagonal es

$$\sum_{b} \sum_{c>b} x_{bc} = 3 + 1 + 2 + 1 = 7$$

y las discrepancias observadas, expresadas como promedio de todos los pares posibles, x.., son

$$D_o = \sum_{b} \sum_{c} \frac{x_{bc}}{x_{..}} d_{bc} = \frac{3+1+2+3+1+1+1+2}{54} = \frac{14}{54} = 0{,}259.$$

Los productos de las entradas marginales en uno de los triángulos que están fuera de la diagonal, que son iguales a x.. multiplicado por la cantidad de pares no coincidentes previstos, son:

$$\sum_{b} \sum_{c>b} x_{b.}\, x_{.c} = 12 \cdot 24 + 12 \cdot 8 + 12 \cdot 10 + 24 \cdot 8 + 24 \cdot 10 + 8 \cdot 10 = 1036$$

que es igual a la suma, aplicada a las entradas del triángulo inferior fuera de la diagonal, de la siguiente matriz de (x.. - m + 1) = 52 por las frecuencias previstas:

12·10	24·12	8·12	10·12	12·52
12·24	24·22	8·24	10·24	24·52
12·8	24·8	8·6	10·8	8·52
12·10	24·10	8·10	10·8	12·52
12·52	24·52	8·52	10·52	54·52

y la discrepancia prevista resulta

$$D_c = \sum_b \sum_c \frac{x_{b.}\, x_{.c}}{x_{..}\,(x_{..} - m + 1)}\, d_{bc} = \frac{12\cdot 24 + 12\cdot 8 + 12\cdot 10 + 24\cdot 12 + \ldots}{54\,(54-3+1)} = \frac{2032}{54\cdot 52} = 0,724$$

de modo tal que el coeficiente de acuerdo, calculado de ambas maneras, pasa a ser

$$\alpha = 1 - \frac{D_o}{D_c} = 1 - \frac{.259}{.724} = 0,642$$

$$\alpha = 1 - (x_{..} - m + 1)\,\frac{\sum_b \sum_{c>b} x_{bc}\; x_{bc}}{\sum_b \sum_{c>b} x_{b.}\, x_{.c}\, d_{bc}} = 1 - (54-3+1)\,\frac{7}{1016} = 0,642$$

Las matrices de coincidencia no sólo resumen los datos sobre fiabilidad con vistas a una más fácil identificación de los errores de codificación, sino que además ponen de relieve con mayor claridad el significado de las diferencias entre categorías y valores. Consideremos las diferencias como una ponderación asignada a cada casillero de una matriz de coincidencia, cuyas categorías abarcaran de 0 a 7, por ejemplo (véase pág. 213).

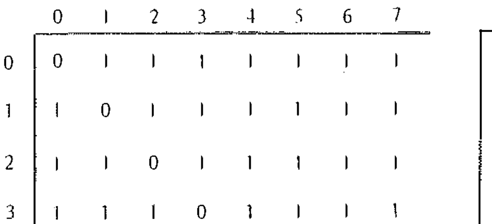
En el caso nominal, todas las entradas que están fuera de la diagonal presentan idénticas diferencias y, por lo tanto, las dos medidas de discrepancia no hacen más que enumerar las ocurrencias observadas y previstas en esos casilleros. Esto ya se ha ilustrado en los ejemplos mencionados anteriormente. Las diferencias de intervalos dependen de la distancia que hay entre las categorías, y en la medida en que las escalas de intervalos no reconocen ningún punto de referencia, no importa entre qué valores se observa o se prevé una diferencia. Las líneas de diferencias iguales corren paralelas, pues, a la diagonal. Por otra parte, las distancias de

razones ponderan pares de valores que son mayores cuanto más próximos se encuentran al punto del cero absoluto. Aquí cabe decir que es la desviación angular respecto de la diagonal la que indica la magnitud de la diferencia entre los dos valores. Consecuentemente, la diferencia entre "nada" y "algo" es en este caso igual a la diferencia entre "nada" y "muchísimo"; no obstante, en el caso de que a alguien se le estafe un dólar, este último tendrá mucho más valor si esa persona sólo tiene dos dólares, que si es un millonario. Y esto es lo que explica la diferencia de razones. La diferencia entre los rangos de las escalas ordinales es más difícil de describir, ya que no depende del valor numérico de sus categorías sino de la cantidad de rangos que se intercalan entre ellos. Supóngase que a diez unidades se les asignan los rangos siguientes: 1, 2, 3, 3, 4, 4, 4, 4, 5, 10. Intuitivamente, la diferencia entre 1 y 3 debe ser mucho menor que la que existe entre 3 y 5, ya que en este caso se puede reconocer que hay muchos más rangos intermedios. Por igual motivo, la diferencia entre 1 y 2 ha de ser la misma que la que hay entre 5 y 10. Aunque la diferencia numérica sea mayor, la cantidad de rangos intermedios es la misma. Aplicando la definición de diferencias ordinales a las frecuencias que acabamos de dar, obtenemos distancias que son iguales al doble de la cantidad de rangos entre el valor central de cada clase:

	1	2	3	4	5	10	
1	0	2	5	11	15	18	$n_1 = 1$
2	2	0	3	9	14	16	$n_2 = 1$
3	5	3	0	6	11	13	$n_3 = 2$
4	11	9	6	0	5	7	$n_4 = 4$
5	16	14	11	5	0	2	$n_5 = 1$
10	18	16	13	7	2	0	$n_{10} = 1$

d_{bc} ordinal (ejemplificado)

Unicamente en el caso de que las frecuencias de los rangos estén uniformemente distribuidas, la diferencia ordinal será de hecho igual a la diferencia de intervalos.



	0	1	2	3	4	5	6	7
0	0	1	1	1	1	1	1	1
1	1	0	1	1	1	1	1	1
2	1	1	0	1	1	1	1	1
3	1	1	1	0	1	1	1	1
4	1	1	1	1	0	1	1	1
5	1	1	1	1	1	0	1	1
6	1	1	1	1	1	1	0	1
7	1	1	1	1	1	1	1	0

d_{bc} nominal

0	1
0	1
1	0
4	1
9	4
16	9
25	16
36	25
49	36

2	3	4	5	6	7
4	9	16	25	36	49
1	4	9	16	25	36
0	1	4	9	16	25
1	0	1	4	9	16
4	1	0	1	4	9
9	4	1	0	1	4
16	9	4	1	0	1
25	16	9	4	1	0

d_{bc} de intervalos

0	1	2	3	4	5	6	7
0	1	1	1	1	1	1	1
1	0	$\frac{1}{3}$	$\frac{2}{4}$	$\frac{3}{5}$	$\frac{4}{6}$	$\frac{5}{7}$	$\frac{6}{8}$
1	$\frac{1}{3}$	0	$\frac{1}{5}$	$\frac{2}{6}$	$\frac{3}{7}$	$\frac{4}{8}$	$\frac{5}{9}$
1	$\frac{2}{4}$	$\frac{1}{5}$	0	$\frac{1}{7}$	$\frac{2}{8}$	$\frac{3}{9}$	$\frac{4}{10}$
1	$\frac{3}{5}$	$\frac{2}{6}$	$\frac{1}{7}$	0	$\frac{1}{9}$	$\frac{2}{10}$	$\frac{3}{11}$
1	$\frac{4}{6}$	$\frac{3}{7}$	$\frac{2}{8}$	$\frac{1}{9}$	0	$\frac{1}{11}$	$\frac{2}{12}$
1	$\frac{5}{7}$	$\frac{4}{8}$	$\frac{3}{9}$	$\frac{2}{10}$	$\frac{1}{11}$	0	$\frac{1}{13}$
1	$\frac{6}{8}$	$\frac{5}{9}$	$\frac{4}{10}$	$\frac{3}{11}$	$\frac{2}{12}$	$\frac{1}{13}$	0

d_{bc} de razones

Supóngase que los valores del ejemplo constituyeran en realidad una escala de intervalos, y que las diferencias (b - c) fueran consiguientemente significativas. En tal caso el coeficiente tendría que ponderar una confusión 1-3 cuatro veces más que una confusión 1-2, y una confusión 1-4 nueve veces más. Presumiendo que sean diferencias de intervalos, y no las diferencias nominales antes supuestas, tenemos

$$\alpha = 1 - \frac{\sum_b \sum_{c>b} \frac{x_{bc}}{x_{..}} d_{bc}}{\sum_b \sum_{c>b} \frac{x_{b.}\, x_{.c}}{x_{..}(x_{..}-m+1)} d_{bc}} = 1 - \frac{\frac{3\cdot 1^2+1\cdot 2^2+2\cdot 3^2+1\cdot 1^2+\cdots}{54}}{\frac{12\cdot 24\cdot 1^2+12\cdot 8\cdot 2^2+12\cdot 10\cdot 3^2+\cdots}{54(54-3+1)}} = 0,547.$$

Si las faltas de coincidencia estuvieran más próximas a la diagonal, como sería de esperar en una "auténtica" escala de intervalos, el valor del coeficiente de intervalos sería superior al del coeficiente que parte del supuesto de una escala nominal. La falta de coincidencia 1-4 reduce el coeficiente de intervalos por debajo del valor antes obtenido. Un examen de la matriz sugiere que las premisas correspondientes a una escala de razones serían aún peores, y de hecho el coeficiente de razones es igual a 0,483.

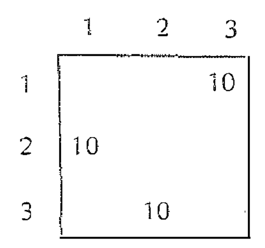
Las matrices de coincidencia entre sólo dos observadores son mucho más fáciles de interpretar. Pueden obtenerse añadiendo simplemente la matriz de contingencia entre dos observadores a su transpuesta; en el ejemplo dicotómico que dimos anteriormente,

$$\text{Jin}\ \overset{\text{Han}}{\begin{bmatrix} 5 & 3 \\ 1 & 1 \end{bmatrix}} + \text{Han}\ \overset{\text{Jin}}{\begin{bmatrix} 5 & 1 \\ 3 & 1 \end{bmatrix}} = \begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix}$$

donde las sumas marginales son iguales al número total de ocurrencias de una categoría $x_b = n_b$. Esto simplifica la conceptualización y cómputo del coeficiente. Si además el número de categorías es igual a 2 (una distinción dicotómica entre 0 y 1), el cálculo del coeficiente toma su forma más sencilla:

$$\alpha = 1 - (x_{..} - 1)\ \frac{x_{01}}{x_{0.}\ x_{.1}}$$

Un error corriente, contra el que conviene estar advertido en este caso, es la confusión entre acuerdo y asociación. La siguiente tabla de contingencia muestra una alta asociación y un acuerdo nulo:



	1	2	3
1			10
2	10		
3		10	

Particularmente engañosa es la costumbre de utilizar coeficientes de correlación para las evaluaciones de la fiabilidad. Si todos los datos pueden representarse en la recta $X = aY + b$, la correlación es total, pero el acuerdo exige que $X = Y$, que no es lo que miden los coeficientes de correlación. Ni las medidas de asociación ni los coeficientes de correlación registran los errores sistemáticos de codificación, y por lo tanto no resultan adecuados como medida de la fiabilidad.

Fiabilidad de los datos y normas

La finalidad última de verificar la fiabilidad es establecer si los datos obtenidos en el curso de una investigación ofrecen una base segura para formular inferencias, hacer recomendaciones, apoyar decisiones o aceptar algo como un hecho. De acuerdo con esta finalidad última, la fiabilidad es un atributo de todas las distinciones relevantes que se establecen dentro de los datos, y se denomina "fiabilidad de los datos". Esta se opone, pues, a otras características de un proceso cuya fiabilidad puede evaluarse principalmente con fines diagnósticos.

Es importante advertir que la fiabilidad de los datos es una cualidad de éstos, y no de la realidad que le interesa al investigador. Esta distinción es importante para el muestreo. Dado que los diseños de fiabilidad exigen una duplicación de los esfuerzos de codificación y de procesamiento de los datos, el hecho de generar una adecuada fiabilidad de los datos duplica, como mínimo, el costo de generar los datos sobre los cuales se han de basar los

hallazgos de la investigación. Por lo tanto, habitualmente se someten a una prueba de fiabilidad no todos los datos, sino una muestra de ellos. Esta submuestra no necesita ser representativa de las características de la población (que puede servir de base para el muestreo de los datos con vistas al análisis), pero sí *debe ser representativa de todas las distinciones establecidas* dentro de la muestra de los datos disponibles. Entonces, muestrear una población con fines de análisis, y efectuar el submuestreo de una muestra de los datos para verificar la fiabilidad, son dos cosas distintas. Para esto último, lo mejor es un diseño de muestreo estratificado que se ocupe de que todas las categorías de análisis, todas las decisiones especificadas por diversos tipos de instrucciones, estén de hecho representadas en los datos sobre fiabilidad, independientemente de la frecuencia con que aparezcan en los datos reales.

Por lo tanto, al verificar la fiabilidad de los datos, uno de los interrogantes a que debe responderse es si los datos sobre fiabilidad son lo bastante representativos de las distinciones establecidas dentro de los datos que se han de analizar. Otro es cuán alto ha de ser el grado de acuerdo para que los datos sean considerados lo suficientemente fiables como para que esté justificado el análisis. De ello nos ocuparemos ahora.

Para empezar, en un análisis de contenido que abarque numerosas variables, la fiabilidad de los datos no puede ser una cifra única. Hay buenos motivos para suponer que lo anterior define el acuerdo sólo para una variable única. Aunque podría calcularse un acuerdo medio para todas las variables que abarca un estudio, ésta sería una cifra engañosa, porque daría por sentado que cualquier variable fiable sería capaz de compensar a cualquier otra no fiable. Pero esto no es posible, en particular si las variables están lógicamente diferenciadas y libres de redundancia conceptual, que es lo que procuran muchos lenguajes de datos utilizados en el análisis de contenido. Si se necesita una enunciación sumaria, *la menor medida de acuerdo de la serie será el mejor indicador de la fiabilidad de los datos.* Cualquier variable podría llegar a convertirse en un obstáculo para la seguridad que ofrecen los datos en su conjunto.

Hemos abogado por la necesidad de tomar el acuerdo menor como medida de fiabilidad de datos multivariados; pero sigue en pie el siguiente interrogante: ¿cuán alto debe ser ese nivel de acuerdo? La fiabilidad de los datos requerida, ¿debe ser como

mínimo 0,95, 0,90 ó 0,80? A pesar de que todos los análisis de contenido afrontan estas preguntas, no existe una respuesta predeterminada.

Con el fin de aclarar de qué manera pueden interpretarse los diferentes niveles de fiabilidad, un colega holandés, Marten Brouwer, creó en cierta oportunidad un conjunto de categorías, les asignó complicados nombres holandeses sin similitud alguna con las palabras inglesas y solicitó a los codificadores que describieran en esos términos a los personajes televisivos. Aunque esas palabras connotaban, en el mejor de los casos, algunas vagas asociaciones entre rasgos morfofonémicos y rasgos de la personalidad, la medida de acuerdo fue ya de 0,44. A partir de las categorías así registradas, nadie que esté en su sano juicio puede justificar que se formulen enunciados acerca de lo que estos codificadores observaron en realidad. En varios análisis de contenido comprobé que las correlaciones entre las variables con un nivel de acuerdo menor que 0,7 tendían a carecer de significación estadística. Por supuesto, esto es bastante obvio, ya que están excesivamente contaminadas por el ruido. En el estudio de BROUWER y otros (1969) adoptamos como política informar acerca de las variables sólo en los casos en que su fiabilidad superara el 0,8, y únicamente admitimos variables cuya fiabilidad oscilaba entre 0,67 y 0,8 para establecer conclusiones sumamente provisionales y cautelosas. Estas normas siguieron vigentes en los trabajos sobre los indicadores culturales (GERBNER y otros, 1979), y podrían servir de guía en otros casos.

En lo posible, *las normas referentes a la fiabilidad de los datos* no deben adoptarse ad hoc, sino que *deben estar relacionadas con los requisitos impuestos en materia de validez a los resultados de la investigación, específicamente a los costos que acarrea extraer conclusiones equivocadas.* Si se tratara de una cuestión de vida o muerte, un analista de contenido ni siquiera debería aceptar una norma sobre fiabilidad de los datos que llevara a un error en los resultados con una probabilidad, digamos, inferior a la de morir en un accidente automovilístico (que parece ser una probabilidad que los seres humanos están dispuestos a aceptar en su vida corriente). Si se trata de un estudio exploratorio, sin consecuencias graves, ese nivel de rigurosidad puede disminuir considerablemente, pero no tanto como para que los hallazgos ya no puedan tomarse en serio.

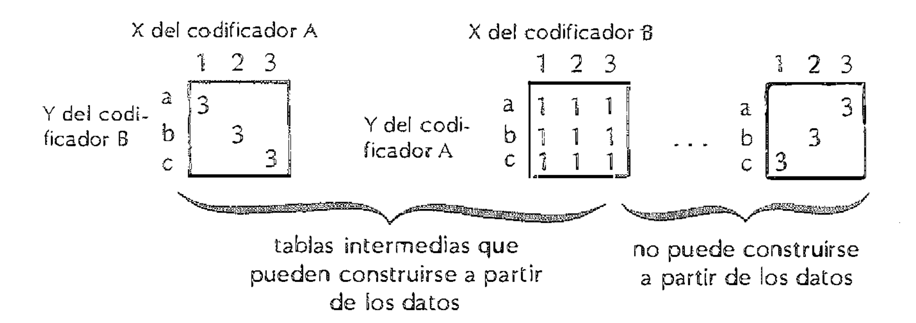
Con el fin de establecer un nivel significativo de fiabilidad, el

analista de contenido tendrá que determinar de qué manera la falta de fiabilidad que aparece en los datos afectará sus hallazgos. Algunos análisis de contenido son muy sólidos, en el sentido de que la falta de fiabilidad apenas se nota en el resultado. En otras situaciones, un pequeño número de errores puede generar enormes diferencias, convirtiendo una respuesta afirmativa en su opuesta. Para verificar la sensibilidad de un análisis de contenido respecto de la falta de fiabilidad en los datos, debe analizarse y procesar no sólo un simple conjunto de datos sino tantos como puedan obtenerse por permutación de las categorías entre las cuales se encontró la discrepancia. *La falta de fiabilidad en los datos define una distribución de resultados posibles* entre los cuales es probable que aparezca el hallazgo "verdadero". Si esta distribución es más amplia de lo que puede tolerarse, la fiabilidad de los datos es excesivamente baja.

Para demostrar lo que queremos decir con esto, supóngase que interesa establecer si dos variables están asociadas entre sí, y supóngase además que ambas son registradas por dos codificadores independientes:

	unidades	1	2	3	4	5	6	7	8	9
variable X	codificador A	1	2	3	1	2	3	1	2	3
	codificador B	1	2	3	2	3	1	1	2	3
variable Y	codificador A	b	c	a	a	b	c	a	b	c
	codificador B	a	b	c	a	b	c	a	b	c

La fiabilidad para ambas variables es 0,628. Puesto que es imposible saber si puede confiarse en el codificador 1 o en el codificador 2, o en qué momento puede confiarse en ellos, por el simple procedimiento de elegir para cada unidad ya sea de la primera o de la segunda columna, pueden construirse toda una gama de pautas de posibles co-ocurrencias a partir de estos datos:



Si no se controlaran los efectos de la variabilidad del error que introduce la falta de fiabilidad, podría tomarse como resultado cualquiera de estas pautas. De acuerdo con este razonamiento, y dado que alrededor de la mitad de las posibles co-ocurrencias no pueden construirse a partir de los datos, no es probable que las asociaciones sean negativas; sin embargo, nada puede decirse de hecho en cuanto a si existe una asociación positiva, o en cuanto a su magnitud. Unicamente un nivel más alto de acuerdo entre los codificadores puede estrechar la amplia distribución de pautas posibles. Si el objetivo es establecer asociaciones, 0,628 no sería un nivel de acuerdo aceptable.

Un error corriente en la bibliografía sobre análisis de contenido consiste en verificar la hipótesis nula sobre los datos acerca de la fiabilidad y aceptar que éstos son fiables si dicha hipótesis resulta insostenible. Este error es grave. Al verificar la fiabilidad de los datos, lo que uno desea es asegurarse de que los errores se mantienen dentro de límites tolerables o, lo que es lo mismo, que el nivel de acuerdo alcanzado se aparta sólo mínimamente de una correspondencia total. Un acuerdo que simplemente se aparte del azar, por significativo que sea este desvío, no ofrece seguridad alguna de que pueda confiarse en los datos. Como no parece posible verificar la significación estadística del desvío respecto del acuerdo total, lo que se necesita es, ante todo, que todos los acuerdos sean superiores al que requiere un análisis particular; y en segundo lugar, si lo son, se precisa verificar la hipótesis de que podría habérselos obtenido a partir de una población de datos sobre fiabilidad cuyo acuerdo fuese inferior al requerido. En caso de que no se cumpla esta hipótesis, los datos son verdaderamente fiables en un grado superior al previsto.

PROCEDIMIENTOS DE DIAGNOSTICO

Mientras se va creando el instrumento para el análisis de contenido, y mucho antes de verificar la fiabilidad de los datos, pueden llevarse a cabo numerosas pruebas de fiabilidad preliminares y detalladas a fin de identificar las fuentes de la falta de fiabilidad en el diseño de investigación. El diagnóstico de los problemas que presenta este último puede dar origen a modificaciones de las categorías del análisis, de las instrucciones proporcionadas a los codificadores e incluso de las decisiones acerca de qué personas participarán en el proceso, hasta que se eliminen todos los obstáculos. Es muy habitual que se efectúen dos o tres pruebas diagnósticas de fiabilidad antes de que el investigador se sienta seguro como para seguir adelante, verificando la fiabilidad de los datos y generando el resto de los datos para el análisis.

En lo que sigue, distinguiremos y ejemplificaremos cuatro procedimientos de diagnóstico:

- fiabilidad de la unidad
- fiabilidad de los individuos participantes
- fiabilidad de una categoría única
- fiabilidad condicional

El primer procedimiento identifica los problemas de falta de fiabilidad en el material que sirve de puente; el segundo, en los codificadores participantes; y los dos últimos, en las instrucciones de registro.

Fiabilidad de la unidad

La *fiabilidad de la unidad* sitúa la fuente de la falta de fiabilidad en las unidades registradas. Algunas de éstas son de más difícil descripción que otras, y las que provocan tropiezos en la codificación necesitan ser examinadas con cuidado con el fin de encontrar el modo de ajustar las instrucciones de registro a las propiedades que aparecen en los datos. En el ejemplo anterior,

		1	2	3	4	5	6	7	8	9
	A	1	1	2	4	1	2	1	3	2
codificador:	B	1	2	2	4	4	2	2	3	2
	C	1	2	2	4	4	2	3	3	2

(unidades: 1–9)

En las unidades 1, 3, 4, 6, 8 y 9 no se aprecian faltas de coincidencias; para estas unidades, la discrepancia observada es 0, y su fiabilidad resulta igual a 1. En las unidades 2 y 5, dos de los seis pares posibles presentan falta de coincidencia, y la discrepancia observada es 0,333 en cada una de las unidades, lo cual da una fiabilidad de 0,539 para ellas. Para la unidad 7, todos los pares presentan falta de coincidencia, el desacuerdo observado es 1 y la fiabilidad pasa a ser -0,382 lo cual es inferior al azar.

Como procedimiento de diagnóstico, la fiabilidad de la unidad no debe utilizarse como criterio para la inclusión o exclusión de determinadas unidades de análisis (ya nos hemos referido a esto anteriormente). No obstante, es admisible utilizarla como criterio para limitar el uso de la instrucción de registro a una *clase* de unidades de registro definida por criterios que no sean el de la fiabilidad. Por ejemplo, en un estudio sobre la programación televisiva, comprobamos que los dibujos animados constituían las unidades menos fiables, lo cual se pudo explicar después de calcular la fiabilidad para los dibujos animados y para otros programas de forma separada: resultó que nuestro instrumento de registro era inadecuado para la categorización de los personajes de los dibujos animados. Enfrentados con la opción de generalizar la instrucción para incluir a estos personajes o circunscribir el estudio, nos decidimos por esta última. Así quedaban restringidas las generalizaciones que podíamos formular, pero no se desviaban los hallazgos en la dirección de las unidades fácilmente codificables.

Fiabilidad de los individuos participantes

La *fiabilidad de los individuos participantes* mide el grado en que alguno de los individuos se erige en fuente de datos poco fiables. Con suma frecuencia, hay codificadores más cuidadosos que otros, o que operan de manera más congruente, o que entienden las instrucciones y el lenguaje fuente mejor que los demás. Allí donde importe la fiabilidad, sólo deberá recurrirse a los más capaces. La fiabilidad de los individuos participantes suministra un medio de determinar el grado en que puede considerarse fiable cada observador, codificador o juez.

Pueden obtenerse datos sobre la fiabilidad de los individuos en dos situaciones distintas. En condiciones de test-norma, ésta pasa a ser una medida de exactitud, que es lo que uno desearía que fuese. Sólo en muy pocos casos (como en la situación de capacita-

ción antes mencionada) se presentan estas condiciones. En condiciones de test-test, el trabajo de un individuo puede compararse con el de los restantes tomados como grupo; son necesarios, pues, más de dos codificadores, preferiblemente muchos. El trabajo del grupo con respecto al cual se compara el de ese individuo, pasa a ser la norma, y la fiabilidad del individuo evalúa hasta qué punto éste se encuentra próximo a la medida central del grupo.

En los datos sobre fiabilidad del ejemplo, el codificador A, con su aparente preferencia por la categoría 1, es a todas luces anómalo. Para construir la matriz de coincidencia de las co-ocurrencias que sólo implican a A, suprimimos de la matriz de coincidencia de todos los pares de co-ocurrencias las que se dan entre B y C y no contribuyen a la comparación:

A:B:C

6	3	1	2
3	20	1	
1	1	6	
2			8

$D_o = \frac{14}{54} = 0{,}259$

$D_e = \frac{2032}{2808} = 0{,}724$

$\alpha = 0{,}642$

B:C

2			
	8	1	
	1	4	
			4

$D_o = \frac{2}{18} = 0{,}111$

$D_e = 0{,}724$

$\alpha = 0{,}846$

= A:B+C

4	3	1	2
3	12		
1		4	
2			4

$D_o = \frac{12}{38} = 0{,}316$

$D_e = 0{,}724$

$\alpha = 0{,}564$

Para los tres codificadores, hay 14 faltas de coincidencias sobre un total de 54 pares posibles; para los codificadores B y C, la relación es de 2/18, mientras que para A y B + C, es de 12/38. Antes calculamos que la discrepancia prevista para los tres codificadores era de 0,724, cifra que debe considerarse la estimación más segura de la situación correspondiente al azar en los tres casos. Arriba se exponen los tres coeficientes. La conclusión a que conduce este análisis es que el codificador A es, en realidad, poco fiable. Si se pudieran eliminar, controlar o recodificar por otros dos codificadores los datos que aporta al estudio, la fiabilidad de los datos mejoraría de 0,642 a 0,846.

Fiabilidad de una categoría única

La *fiabilidad de una categoría única* evalúa el grado en que una categoría se confunde con el resto del conjunto de categorías. Habitualmente, las confusiones de esta índole derivan de una definición ambigua o son consecuencia de la incapacidad del codificador para comprender el significado de dicho valor o categoría. De todas las distinciones que se establecen dentro de una variable, la fiabilidad de una categoría única sólo concierne a las que abarcan la variable en cuestión. En la matriz de coincidencia de nuestro ejemplo, puede verse que la mayor parte de las discrepancias corresponden a la categoría uno. Para eliminar las distinciones entre las categorías restantes, se construye una matriz de coincidencia 2 x 2 sumando las entradas de esta matriz de modo que sólo se mantenga la distinción entre una categoría y todas las demás. Lo que sigue muestra esto para las categorías 1 y 2 por separado:

	1	2	3	4
1	6	3	1	2
2	3	20	1	
3	1	1	6	
4	2			8

$\alpha = 0,642$

	1	otro
1	6	6
otro	6	36

$$D_o = \frac{12}{54} = 0,222$$

$$D_e = \frac{1008}{2808} = 0,358$$

$\alpha = 0,381$

	2	otro
2	20	4
otro	4	20

$$D_o = \frac{8}{54} = 0,148$$

$$D_e = \frac{1152}{2808} = 0,256$$

$\alpha = 0,711$

Los coeficientes que aquí aparecen cuantifican lo que es intuitivamente obvio: la categoría 1 está mal definida y es poco fiable, mientras que la categoría 2 tiene una mejor definición.

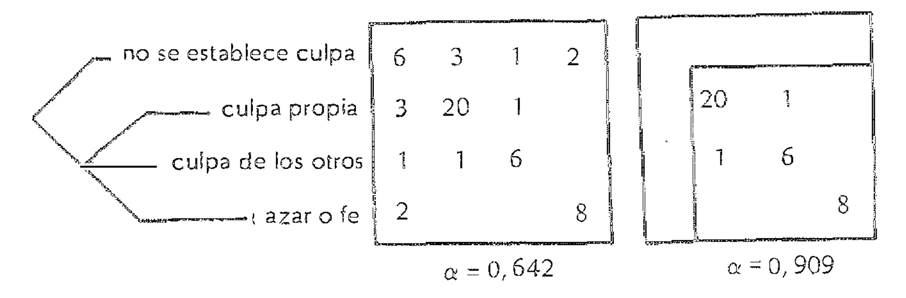
La fiabilidad de una categoría única presupone que las variables no están ordenadas. Si es baja, el investigador dispone de dos alternativas: o bien redefine la categoría poco fiable, aclarando mejor la diferencia entre ésta y las demás, y luego realiza otro test de fiabilidad para ver si la modificación fue en realidad fructífera, o bien no usa la distinción en el análisis subsiguiente y se centra,

en cambio, en las categorías restantes; en este caso tendrá que verificar la fiabilidad de forma condicional.

Fiabilidad condicional

La *fiabilidad condicional* estima la fiabilidad de un subconjunto de categorías dentro de una variable. Es preciso recurrir a ella cuando algunas categorías son tan poco fiables que no es posible basar ninguna conclusión en ellas, aunque las restantes categorías sean aún rescatables. Dichas situaciones son muy corrientes en el análisis de contenido. Por ejemplo, suele ser mucho más arduo obtener un acuerdo acerca de si una categoría determinada está o no presente en alguna unidad, que acerca de la manera de clasificarla una vez que se ha comprobado su presencia. Esto puede aplicarse particularmente a la identificación de rasgos de la personalidad, temas, estados de ánimo, expresiones emocionales, etcétera.

Supóngase que las categorías de nuestro ejemplo pudieran interpretarse como resultado de dos decisiones, siendo el resultado de la primera una condición para que se aplique la segunda. En tal caso, la fiabilidad condicional elimina o desestima el efecto de la primera decisión.



Para obtener las fiabilidades condicionales, sólo se promedian las discrepancias observadas y previstas dentro de la submatriz que contiene las distinciones relevantes, con lo cual las probabilidades de la discrepancia prevista tienen en cuenta todas las distinciones establecidas. En el ejemplo, hay 36 acuerdos respecto de las categorías distintas de 1, dos de las cuales se observa que no coinciden para 2, 3 y 4, de modo que la discrepancia observada es 2/36. La forma en que surge la discrepancia prevista se aprecia en las siguientes matrices: a) la de las frecuencias previstas para la

matriz total, que fueron utilizadas anteriormente para calcular la fiabilidad de los datos, etc.; b) la de las frecuencias previstas para la primera decisión, en la cual el acuerdo previsto respecto de las categorías distintas de 1 equivale a 32,31 entradas; c) la de las probabilidades previstas en la submatriz condicional sobre el acuerdo respecto de las categorías distintas de 1. Para obtener estas probabilidades, las frecuencias previstas de los casilleros de la submatriz se dividen por la frecuencia total de esta submatriz. La contribución de los casilleros de esta submatriz a la discrepancia prevista aparece en las entradas que no están en la diagonal, y suman 64/105. La fiabilidad condicional es 0,909.

(a)

	1	2	3	4
1	120/52	288/52	96/52	120/52
2	288/52	528/52	192/52	240/52
3	96/52	192/52	48/52	80/52
4	120/52	240/52	80/52	80/52

(b)

	1	otro
1	120/52	504/52
otro	504/52	1680/52

(c)

	2	3	4
2	33/105	12/105	15/105
3	12/105	3/105	5/105
4	15/105	5/105	5/105

$D_e = \frac{64}{105} = 0{,}610$

La fiabilidad de una categoría única y la fiabilidad condicional pueden considerarse como las dos partes complementarias de lo que en otro lugar (KRIPPENDORFF, 1971) hemos denominado "fiabilidad de la decisión". Esta verificación permite al analista examinar la fiabilidad de cada división en un árbol de decisiones complejas, suprimiendo los efectos de las decisiones previas en la forma aquí descrita para la fiabilidad condicional, y pasando por alto las discrepancias de las decisiones posteriores, según se ejemplificó en la fiabilidad de una categoría única. La figura 21 muestra un ejemplo de un árbol de decisiones de esa índole. Para evaluar la fiabilidad de la decisión sobre los tipos de valor, se eliminan, mediante adición, las distinciones introducidas por todas las decisiones subsiguientes, y se considera que el resultado de la decisión anterior es la condición necesaria para formular juicios sobre los tipos de valor.

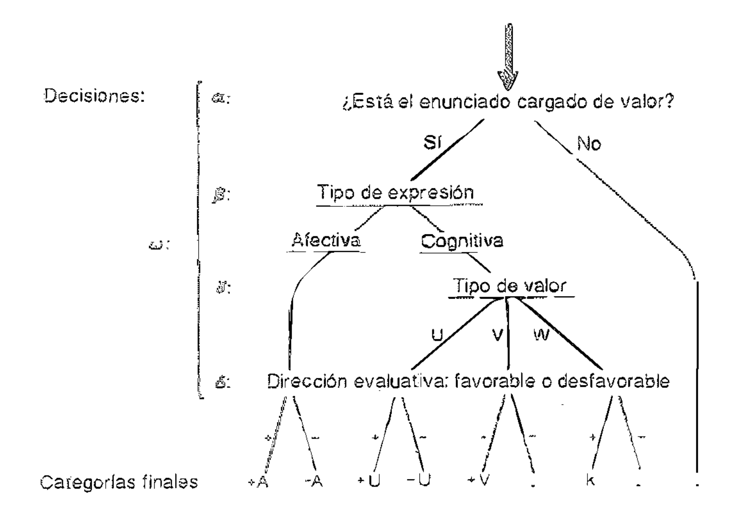


Figura 21. Proceso ramificado de decisiones cuya fiabilidad se evalúa por separado.

Perfeccionamiento de la fiabilidad

La fiabilidad condicional es uno de los procedimientos para determinar qué partes de una variable, cuya fiabilidad total es tan baja que resulta inaceptable, podrían mantenerse en el análisis final. Hay otros procedimientos, todos los cuales en definitiva verifican si una u otra transformación de los datos será capaz de mejorar la fiabilidad.

Uno de ellos consiste en verificar los efectos de *agrupar categorías*. Por ejemplo, en la siguiente matriz de coincidencia aparecen confundidas las categorías 2 y 4, que los dos codificadores no pueden distinguir de manera fiable. Si el investigador puede permitirse dejar a un lado esta distinción y agrupar las dos categorías en una sola, esta variable modificada bien puede volverse lo bastante fiable como para resultar admisible en el análisis.

	1	2	3	4	5
1	10				
2		6	1	6	
3		1	8		
4		6		6	
5					12

$\alpha = 0,692$

	1	2&4	3	5
1	10			
2&4		24	1	
3		1	8	
5				12

$\alpha = 0,950$

Huelga decir que este agrupamiento de categorías exige realizar ciertos supuestos sobre las escalas nominales. Pero la decisión de establecer distinciones más groseras dentro de una variable ordenada (por ejemplo, modificar la codificación de la edad de una persona, haciéndolo en amplias categorías de edad social, en lugar de hacerlo en años) tiene el mismo efecto, a la vez que mantiene el orden.

Por último, en ciertas oportunidades, los errores de codificación son sistemáticos. Por ejemplo, un codificador puede situar un cierto concepto en la categoría X, mientras que otro lo hace permanentemente, y sin ninguna mala intención, en la categoría Y. Si este error es sistemático, habrá más co-ocurrencias X-Y o Y-X que co-ocurrencias X-X o Y-Y. El acuerdo puede volverse negativo (lo que siempre es una señal de error sistemático). Estos errores pueden identificarse en una *matriz de contingencia*, y posteriormente eliminarse encontrando una explicación para ellos. (Si los errores son aleatorios, no se encontrará dicha explicación.) Análogamente, cuando se codifican los conceptos en escalas de diferencial semántico que van de + 3 a -3, algunos codificadores evitan los puntos extremos mientras que otros evitan el punto central. Una transformación de los datos de uno de los codificadores puede rectificar la naturaleza sistemática del desacuerdo y hacer los datos más fiables, siempre y cuando se comprenda cabalmente la índole sistemática de estos errores.

13. Validez

De todos los hallazgos científicos se pretende que sean válidos, en el sentido de representar fenómenos reales. En este capítulo se desarrolla una tipología referida a las clases de validez importantes para el análisis de contenido, y se muestra cuáles son los requisitos necesarios para evaluar cuantitativamente las distintas clases de validez.

La palabra "validez" designa esa propiedad de los resultados de las investigaciones que lleva a aceptarlos como hechos indiscutibles. Otras expresiones próximas a ella son "verdad empírica", "exactitud predictiva" y "congruencia con el saber establecido". Decimos que un instrumento de medición es válido si mide lo que está destinado a medir, y consideramos que un análisis de contenido es válido en la medida en que sus inferencias se sostengan frente a otros datos obtenidos de forma independiente.

La importancia de la convalidación radica en la seguridad que ofrece, pues implica que los hallazgos de las investigaciones previas tienen que tomarse muy en serio al construir teorías científicas o al adoptar decisiones en cuestiones prácticas. Dicha seguridad es particularmente conveniente cuando los resultados de un análisis de contenido, o las teorías a las que dan lugar, pueden tener consecuencias en materia de medidas oficiales, o se pretende contribuir con ellos al desarrollo de algún área del gobierno o de

la industria, o presentarlos como pruebas ante los tribunales, o bien cuando afectan a la vida de los seres humanos. En estas situaciones, unas conclusiones erróneas pueden dar lugar a consecuencias sumamente graves. Cabe sostener que la convalidación reduce el riesgo que implica actuar sobre la base de hallazgos de investigación equívocos como si fueran verdaderos.

En el pasado, los analistas de contenido no se preocuparon de manera sistemática por convalidar sus resultados. Tal vez esto explique que otras técnicas de investigación —principalmente los experimentos controlados y la investigación por encuestas— se prefirieran con frecuencia al análisis de contenido, siempre que se presentara la opción. Existen por lo menos dos obstáculos para la convalidación en el análisis de contenido, uno de orden conceptual y otro de orden metodológico. Nuestro marco de referencia conceptual pretende superar ambos.

El obstáculo conceptual que se alza contra la convalidación proviene fundamentalmente de las incertidumbres relacionadas con el objetivo de extraer inferencias a partir de los datos, y (ésta es la otra cara de la misma moneda) de las ambigüedades relacionadas con el hecho de encontrar pruebas independientes que corroboren los resultados. Supóngase que un analista de contenido sostiene haber descubierto un cambio en las actitudes de los Estados Unidos que va de la exaltación de los valores materiales a otros espirituales, pero afirma que ese cambio se observa exclusivamente en los usos de que son objeto los símbolos. A menos que este analista especifique de antemano la clase de fenómenos observables de forma independiente en los que dicho cambio se pondría de manifiesto (en caso de ser cierto), su hallazgo no puede convalidarse, y sólo puede sostenerse a sí mismo. Incertidumbres de esta índole se introducen en las concepciones sobre el análisis de contenido como "meramente descriptivas de un medio de comunicación". Cabe sospechar que algunos analistas de contenido prosperan gracias a estas incertidumbres conceptuales, ya que les ayudan a eludir los peligros implícitos en la refutación de los resultados, aun a riesgo de hacer afirmaciones empíricamente irrelevantes.

El obstáculo metodológico procede principalmente de una interpretación estrecha de la validez. Considérese el siguiente "trilema": si un analista de contenido carece de todo conocimiento directo acerca de lo que infiere, en realidad nada podrá decir sobre la validez de sus hallazgos; si posee algún conocimiento sobre el

contexto de los datos y lo utiliza en la elaboración de sus construcciones analíticas, ese conocimiento deja de ser independiente de su procedimiento, y por ello no puede utilizarlo para convalidar los hallazgos; y si, en fin, se las ingenia para mantener separado el conocimiento sobre el objetivo de sus inferencias respecto del procedimiento que utiliza, sus esfuerzos para obtener inferencias de los datos serán en rigor superfluos, y a lo sumo añadirán un mero caso incidental a la generalización del procedimiento. No obstante, el "trilema" puede solucionarse si se confía en diversas variantes de *pruebas parciales e indirectas* sobre los fenómenos que interesan, o en *pruebas similares* en su forma pero no idénticas a los hallazgos, dejando la *prueba directa* para los empeños de convalidación *post-facto*, y por consiguiente retroactivas.

El eje principal de este capítulo es el tipo de pruebas indirectas que pueden traerse a colación en el caso de los resultados de los análisis de contenido y el modo en que pueden contribuir a la validez de los hallazgos. Nuestra clasificación se aparta un poco de la propuesta en la bibliografía sobre tests psicológicos, sobre todo porque la situación del analista de contenido es distinta de aquella en la que se encuentra el psicólogo experimental.

TIPOLOGIA DE LOS PROCEDIMIENTOS DE VALIDACION

Siguiendo a CAMPBELL (1957), distinguimos entre validez interna y validez externa. La *validez interna* no es sino otra manera de designar la fiabilidad. Emplea criterios inherentes a un análisis y evalúa si los hallazgos de la investigación tienen alguna relación con los datos disponibles, sin decir cuál. En este caso, sólo nos ocupamos de la *validez externa*, o validez propiamente dicha, que evalúa el grado en que las variaciones inherentes al proceso del análisis se corresponden con las externas a él, y si los hallazgos representan a los fenómenos reales en el contexto de los datos, tal como se pretende. El requisito de que los procedimientos del análisis de contenido sean sensibles al contexto es una exigencia de validez externa. En el caso ideal, todas las etapas de un análisis de contenido deben justificarse utilizando criterios de validez externa, aunque en la práctica eso sólo es posible en determinados puntos críticos.

Aunque varios analistas de contenido se atienen, en lo referente a sus conceptos de validez, a las "Technical Recommendations for

psychological tests and diagnostic techniques" de la AMERICAN PSYCHOLOGICAL ASSOCIATION (1954), aquí nos ocuparemos de los problemas específicos que plantea el análisis de contenido. De acuerdo con nuestra tipología, comenzaremos por distinguir si las pruebas validantes se refieren a la naturaleza de los *datos*, a los *resultados* analíticos o a la índole del *proceso* que conecta aquéllos con éstos.

La *validez orientada a los datos* evalúa hasta qué punto un método de análisis es representativo de la información inherente a los datos disponibles o de la asociada con ésta. Mientras que la fiabilidad asegura que la representación de los datos primitivos en un análisis sea congruente dentro de un mismo método, o entre varios métodos diferentes, o con relación a una norma, la validez orientada a los datos evalúa además el grado en que dicha representación se corresponde con un criterio externo. Aunque tiende a utilizarse para justificar los pasos iniciales de un análisis de contenido, no se limita a éstos. Distinguimos, además, dos subclases de validez orientada a los datos: la validez semántica y la validez del muestreo.

La *validez semántica* evalúa el grado en que un método es sensible a los significados simbólicos relevantes dentro de un contexto determinado. En el análisis de contenido, se logra una alta validez semántica cuando la semántica del lenguaje de los datos se corresponde con la fuente, el receptor o cualquier otro contexto respecto del cual se examinan dichos datos. El interés por las categorías émicas o espontáneas, en lugar de las categorías éticas o impuestas de análisis, refleja también la preocupación por la validez semántica. Esta última, rara vez constituye un problema en las técnicas de investigación en que el proceso de recopilación de los datos está estructurado (véanse las pertinentes distinciones en el capítulo 2, "Fundamentos conceptuales").

La *validez del muestreo* evalúa el grado en el que los datos disponibles son, ora una muestra no tendenciosa de un universo que nos interesa, ora lo suficientemente similares a otra muestra del mismo universo como para considerarlos estadísticamente representativos de éste. Lo ideal es que los datos se muestreen activamente a partir del universo por parte del propio analista; pero, en el análisis de contenido, con suma frecuencia los datos son suministrados por una fuente que utiliza sus propios criterios de selec-

ción, por lo general tendenciosos, criterios que el analista no está en condiciones de controlar. En uno y otro caso, una muestra es válida en la medida en que su composición (ya sea en proporción, escala o distribución) se corresponde con la del universo que se pretende representar con ella.

La *validez orientada a los resultados*, o *validez pragmática*, evalúa hasta qué punto un método "funciona" bien en una variedad de circunstancias. Aquí no hay preocupación alguna por la configuración interna del procedimiento. El éxito global de un análisis de contenido se establece demostrando que sus resultados coinciden o están correlacionados con lo que pretenden representar. También en este caso distinguimos dos clases: la validez correlacional y la validez predictiva.

La *validez correlacional* es el grado en que los hallazgos apoyados en un método guardan correlación con los apoyados en otro, justificando así que se sustituyan entre sí. En este caso, la validez correlacional designa tanto una alta correlación entre las inferencias de un análisis de contenido y otras medidas de similares características contextuales (validez convergente), como la baja correlación entre dichas inferencias y medidas de características distintas (validez discriminante).

La *validez predictiva* es el grado en que las predicciones obtenidas mediante un método concuerdan con los hechos observados de forma directa. En el análisis de contenido, la validez predictiva exige que las inferencias obtenidas muestren un alto grado de acuerdo con los estados, atributos, sucesos o propiedades del contexto de los datos a que esas inferencias se refieren (independientemente de que se trate de fenómenos pasados, contemporáneos o futuros) y un alto grado de desacuerdo con las características contextuales que esas diferencias excluyen lógicamente.

La *validez orientada al proceso* evalúa el grado en que un procedimiento analítico sirve de modelo de ciertas relaciones en el contexto de los datos, las imita o las representa funcionalmente. Dentro del análisis de contenido, esta clase de validez concierne básicamente a la índole de la construcción analítica, que es aceptada o rechazada según la demostrable correspondencia estructural-funcional entre los procesos y categorías de un análisis y las

teorías, modelos y conocimientos aceptados del contexto del que proceden los datos. No encontramos motivos para establecer distinciones sutiles, y por ello llamamos a la validez orientada al proceso, *validez de la construcción*. Así, pues, la tipología propuesta es la que muestra la figura 22.

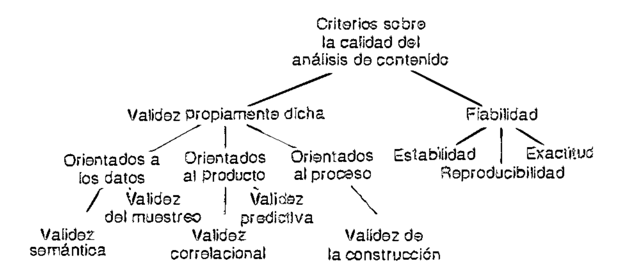


Figura 22. Tipología para la verificación de la validez en el análisis de contenido.

Quizá sea útil detenerse un poco en las similitudes y diferencias entre nuestras distinciones y las propuestas por otros autores. La diferenciación entre la validez predictiva y la correlacional se corresponde con la establecida por JANIS (1965) entre los métodos *directos* y los métodos *indirectos* de validación. Un método directo de validación implica mostrar que los resultados del análisis de contenido describen lo que se pretende que describan, pero como, según Janis, los significados de los mensajes median entre las percepciones y respuestas de los individuos, y en este sentido no son directamente observables, un método indirecto de validación implica "inferir la validez a partir de la productividad" (1965, pág. 70), mientras que "se dice que un procedimiento de análisis de contenido es productivo en la medida en que se comprueba que los resultados que ofrece están correlacionados con otras variables" (1965, pág. 65).

En las "Technical Recommendations" antes mencionadas se distingue entre la validez predictiva y la "validez concurrente", según que una prueba produzca inferencias sobre el futuro o el estado actual de un individuo en lo referente a ciertas variables, medidas en el momento de la prueba. En ambos casos, las "Technical Recommendations" exigen que los hallazgos y variables utilizados como criterio se correlacionen entre sí. En el análisis de

contenido, la dimensión temporal resulta algo secundaria respecto del hecho de que se busca una alta correlación, que indique posibilidad de sustitución más que exactitud predictiva. De ahí que hayamos considerado ambos tipos como validez correlacional.

En las "Technical Recommendations" se habla de una "validez de contenido", que se establece "determinando hasta qué punto... la prueba ofrece una muestra de las clases de situaciones o de materias sobre las cuales se desea extraer conclusiones" (1954, pág. 13). Esto es exactamente equivalente a nuestra definición de la "validez de muestreo". La elección de una expresión distinta está motivada simplemente por la posible confusión que, a nuestro entender, podría introducir el término "contenido".

La expresión "validez pragmática" ha sido tomada de Selltiz y otros, quienes añaden a su definición que "el investigador no necesita saber... *por qué motivo* el rendimiento en la prueba es un indicador eficiente de la característica que le interesa" (SELLTIZ y otros, 1964, pág. 157).

La diferenciación entre validez de la construcción y validez pragmática se ha sostenido también, implícitamente,

> al diferenciar entre los dos tipos de justificación que Feigl (1952) llamó "validación" y "vindicación". En este contexto, la *validación* es una modalidad de justificación según la cual el grado de aceptabilidad de un determinado procedimiento analítico se establece demostrando que es derivable de los principios o teorías generales aceptados, independientemente del procedimiento que se ha de justificar. En cambio, la *vindicación* puede convertir en aceptable un método analítico basándose en el hecho de que conduzca a predicciones exactas (en un grado mayor que el azar), independientemente de los pormenores de dicho método. Para la validación son esenciales las reglas de inducción y deducción, mientras que la base de la vindicación es la relación entre los medios y los fines particulares (KRIPPENDORFF, 1969a, pág. 12).

En las "Technical Recommendations", la validez de la construcción alude a una manifestación no explícita de un rasgo o fenómeno en varias pruebas diferentes, mientras que nuestra definición subraya la correspondencia estructural de un análisis con el conocimiento explícito sobre el sistema que interesa.

Debe advertirse que nuestra tipología no incluye la *validez nominal*, ya que esta forma de justificar los resultados de la investigación no exige ninguna prueba formal, y está totalmente regida

por la intuición. Aunque la intuición no puede omitirse en ninguna etapa de un análisis de contenido, desafía toda elucidación sistemática, que es lo que en este capítulo nos interesa.

El diagrama que se presenta en la figura 23 muestra las clases de comparaciones que se establecen al poner a prueba los cinco tipos de validez mencionados.

VALIDEZ SEMANTICA

Casi todos los análisis de contenido implican reconocer el significado transmitido por las unidades de registro, cierta identificación de las referencias o algunos rasgos semánticos de los datos con que se cuenta. De hecho, para las definiciones antiguas, un requisito del análisis de contenido era "la clasificación de los vehículos-signos" (JANIS, 1965), "la descripción del contenido manifiesto de la comunicación" (BERELSON y LAZARSFELD, 1948, pág. 6), la "codificación" (CARTWRIGHT, 1953, pág. 424), o la "colocación de una variedad de configuraciones de palabras en (las categorías propias de) un esquema clasificatorio" (MILLER, 1951, pág. 96). Aunque muchos análisis de contenido modernos apuntan a inferencias que van mucho más allá de estos procedimientos, no debe subestimarse ni la importancia científica ni la significación práctica de las tareas puramente descriptivas. Aunque las inferencias a las que se apunte dejen muy atrás los significados convencionales, la correcta representación de las cualidades simbólicas de los datos puede ser esencial.

Ya hemos dicho que toda descripción, por simple que sea, debe considerarse una forma de inferencia, y plantearse como interrogante el motivo por el cual diferentes unidades de análisis se sitúan dentro de la misma categoría, así como el criterio o patrón con que eso se hace. El analista puede justificar su descripción remitiéndose a su práctica científica, o a su coincidencia con las diferenciaciones populares, o a las idiosincrasias de una fuente determinada. Sea como fuere, sus descripciones nunca son del todo arbitrarias y puede que deban ser objeto de validación, especialmente si el análisis subsiguiente ya no se justifica en función del contexto de los datos. Es lo que ocurre, por ejemplo, en la mayoría de los análisis de contenido que dan por resultado cómputos de frecuencia, y en los cuales los datos son sometidos a formas de manipulación más sofisticadas, como el análisis factorial.

La operación de computar y de extraer factores a partir de datos multivariados rara vez tiene un equivalente funcional en el mundo social, y por ello la validez de la construcción es a menudo incierta. Pero lo que puede evaluarse en dichas situaciones es si los datos captan o no aquellas cualidades simbólicas que son "reales" de acuerdo con cierto patrón, y si el procedimiento es, por consiguiente, válido hasta el punto de representar dichas cualidades.

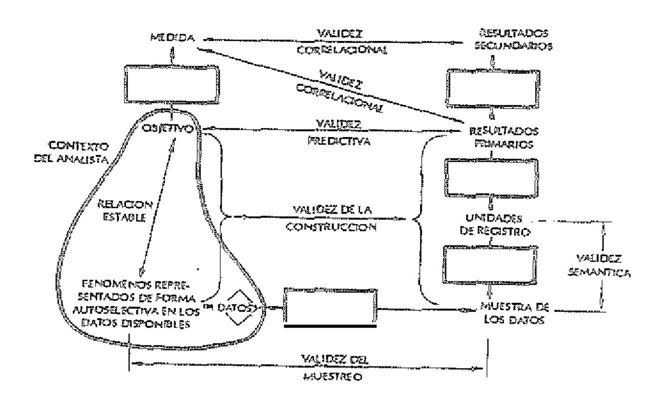


Figura 23. Comparaciones establecidas para diferentes tipos de validez.

Tal vez un análisis de los valores a través de documentos políticos pueda servir como ejemplo muy simple de un trabajo de perfeccionamiento con respecto a la validez semántica. Como primer paso de un elaborado esquema analítico, teníamos que identificar, dentro de un texto determinado, lo que denominamos "enunciados cargados de valor". Cierto es que un grupo de expertos políticos podría haberlos escogido con un alto grado de fiabilidad; pero siguiendo nuestras explícitas instrucciones de registro, los codificadores poseían niveles muy diversos de capacidad, y un programa de ordenadores que habíamos tenido la esperanza de poder emplear resultó en este caso virtualmente inútil. No narraré toda la historia, pero diré que comenzamos por distinguir entre los enunciados que contenían o no ciertos símbolos políticos presentes en cierta lista, entre los que se incluían la democracia, la libertad, la victoria, etcétera, y terminamos por poner a prueba cada enunciado en lo referente a su posible correspondencia con una serie de definiciones estructurales de la forma en que podían expresarse los valores (KRIPPENDORFF, 1970b). Durante la evolución del ins-

trumento de registro, se midió su "perfeccionamiento" por el grado en el que la serie de enunciados identificados como cargados de valor por el instrumento utilizado se encontraban próximos a la serie de enunciados identificados como tales por los expertos. Desde luego, podría cuestionarse el hecho de que el uso de juicios de expertos como criterio de validez semántica se considere significativo, y alguien preguntará por qué no recurrimos a sujetos cualesquiera o a actores políticos; pero lo cierto es que existían limitaciones prácticas, así como motivaciones teóricas, para elegir como referente del análisis este contexto, y nos produjo satisfacción que en última instancia el instrumento escogido se aproximase a lo que afirmaban los expertos en política.

La prueba para la validez semántica de un análisis de contenido es de muy simple elaboración. Toda clasificación de vehículos-signos, toda descripción de unidades de registro, toda codificación de observaciones en un lenguaje descriptivo, divide de hecho la muestra de unidades del análisis en clases de equivalencia mutuamente excluyentes y exhaustivas. Las diferencias que pudieran existir entre las unidades que son asignadas a la misma categoría se pasan por alto.

La validez semántica viene indicada por un grado sustancial de acuerdo entre dos particiones diferentes del mismo conjunto de unidades de análisis. Una partición se obtiene por la manera en que el procedimiento analítico asigna las unidades a sus respectivas categorías; la otra, según un criterio externo, por ejemplo, consiguiendo la participación de ciertos jueces que estén familiarizados con la naturaleza simbólica del mismo lenguaje, y agrupen esas unidades de acuerdo con su similitud semántica. El acuerdo en estas particiones asegura que las unidades pertenecientes a una de las clases de equivalencia comparten las propiedades indicadas por la categoría a las que son asignadas, y que las unidades que pertenecen a clases distintas difieren de acuerdo con las diferencias que existen entre las categorías.

Aunque se hayan emprendido las validaciones semánticas correspondientes, éstas pueden exigir un gran esfuerzo. Antes de proceder a una prueba formal puede recurrirse a dos fáciles controles de validez.

En primer lugar, el analista puede procurarse una lista de todas las unidades agrupadas de acuerdo con las distinciones mantenidas dentro de un estudio. Ante una subdivisión de esa índole, con frecuencia el analista se sorprende al comprobar que su instru-

miento no logra distinguir unidades enormemente distintas entre sí, mientras que por otro lado distingue unidades que en apariencia son semejantes. Dunphy proporciona un ejemplo de todo esto, puesto que obtuvo una tira impresa de la "palabra clave en contexto" para la palabra inglesa "*play*" (véase la figura 20), a fin de demostrar qué acepciones de esa palabra se pasarían por alto en un programa de ordenador que identificara solamente los casos en que aparecía (DUNPHY, 1966, pág. 159). El trabajo demuestra que un análisis que conceda importancia a la distinción entre los diferentes sentidos de la palabra "*play*" no puede permitirse confiar en las meras identificaciones léxicas. Un procedimiento válido desde el punto de vista semántico debe ser sensible a los contextos lingüísticos de la palabra. Dunphy demostró también que aunque un diccionario de palabras de señalamiento puede ser válido con respecto a la asignación de señaladores a palabras únicas, el procedimiento dentro del cual se inserta tal vez no sea adecuado para representar las propiedades simbólicas transmitidas en unidades más amplias de un texto.

En segundo lugar, el analista puede crear un conjunto de unidades hipotéticas de propiedades bien conocidas, y ver de qué manera las distingue su procedimiento. Este método es tradicional en lingüística, donde cuando algún estudioso propone una gramática que pretende distinguir entre oraciones gramaticales y agramaticales, aparece otro con ejemplos contrarios, de oraciones en que la gramática propuesta por el primero no sería capaz de establecer la diferencia. En su crítica al análisis de contingencia, OSGOOD (1959, págs. 73-77) emplea este tipo de razonamiento. En efecto, observa que si el paciente de un psicoanalista enuncia:

1) "Yo amo a mi madre"

un análisis de contingencia añadiría a esta declaración incidental, tal como debe ser, la *asociación* entre el verbo AMAR y el sustantivo MADRE.

Ahora bien, considérense los enunciados siguientes:

2) Siempre amé a mi madre más que a ninguna otra persona.
3) Mi madre me amaba tiernamente.
4) He amado a mi madre.
5) ¿Que yo siempre amé a mi madre? ¡Vaya un chiste!
6) Mi amado padre odiaba a mi madre.

Como las dos palabras críticas co-ocurren en los seis enunciados, un análisis de contingencia los situaría a todos en la misma clase de equivalencia de co-ocurrencia de AMOR-MADRE. No obstante, vemos que, con relación al enunciado 1, en el caso del enunciado 2 el análisis de contingencia es insensible a la intensidad de la asociación expresada; en el caso del enunciado 3, es incapaz de distinguir entre una construcción activa y una pasiva; en el caso del enunciado 4, es insensible a la ausencia actual de un sentimiento que pudo existir antes; en el caso del enunciado 5, es insensible a la ironía, y en el caso del enunciado 6, es insensible a las consideraciones gramaticales.

Así pues, el examen crítico de las distinciones semánticas que establece o que descarta un análisis puede ofrecer valiosas intelecciones acerca de la naturaleza del procedimiento, y quizá suministrar motivos suficientes para rechazar sus resultados.

Validez del muestreo

La validez del muestreo está en juego siempre que la muestra que se estudia difiera del universo que nos interesa, y cuando esta diferencia obedezca a los procesos probabilísticos de selección de las unidades de análisis. Dichos procesos de selección están presentes prácticamente en todos los datos de un análisis de contenido. Debe distinguirse el muestreo realizado por el analista del automuestreo en el interior de la fuente de los datos.

La teoría sobre el *muestreo del analista* ya está suficientemente desarrollada. A partir de un plan de muestreo adecuado, que asegure que cada unidad del universo puede tener las mismas probabilidades de ser incluida en la muestra, la validez del muestreo está virtualmente asegurada para muestras de gran tamaño. No obstante, a medida que las muestras son menores, el error de muestreo reduce dicha validez, hasta un punto en que la semejanza entre la muestra y el universo deja de ser segura.

En lo referente al segundo proceso, las investigaciones sobre comunicación prueban con suma claridad que todos los procesos sociales en los que intervienen comunicaciones y que están representados mediante símbolos o índices son selectivos, son estadísticamente tendenciosos y, de hecho, son el producto de *procesos de automuestreo*. Ténganse en cuenta, por ejemplo, la representación excesiva de personas en edad de casarse y de "anglosajones

blancos y protestantes" de buena posición social entre los personajes televisivos; o la manera selectiva en que los testigos que declaran en los tribunales recuerdan acontecimientos; o la poca cantidad de individuos que nos han permitido conocer tanto la historia como la mitología; o la escasa cantidad de declaraciones públicas sobre las cuestiones relativas a los derechos civiles, comparada con la difusión del prejuicio social y de la injusticia. *Los procesos de automuestreo atribuyen casi siempre probabilidades desiguales* a los sucesos que los individuos, instituciones o culturas consideran merecedores de retención y análisis. Dichos procesos suelen ser bastante sistemáticos, pueden describirse estadísticamente o en términos sociológicos o psicológicos y, por lo tanto, son en alguna medida reconocibles.

Los procesos de automuestreo son anteriores a la disponibilidad de los datos. *El muestreo debe anular las desviaciones estadísticas inherentes a la manera en que el analista accede a los datos*. Y la validez del muestreo procura evaluar el éxito de este trabajo, por lo cual el criterio de validez externa implica conocer las características de automuestreo de la fuente de datos.

Supóngase, por ejemplo, que la tarea consiste en evaluar las percepciones de alguna cuestión política (por ejemplo, las relaciones entre los Estados Unidos y la República Popular China) por parte de los encargados de tomar las decisiones políticas en Vietnam. Una técnica infalible, pero que en esta situación resultaría poco práctica, sería enumerar a los integrantes del círculo vietnamita encargado de tomar las decisiones y entrevistar a una muestra de ellos elegida al azar, asegurándose de que emitieran sus opiniones lo más libremente posible. Un análisis de contenido podría tratar de aproximarse a los resultados de una encuesta sin resultar entrometido. Un analista de contenido ingenuo comenzaría por averiguar con qué datos, informaciones de prensa y otra clase de informes y de crónicas periodísticas se cuenta, y extraería luego una muestra representativa de todos ellos. Aunque podría estratificar dicha muestra, considerando que las cuestiones de política exterior pueden ser objeto de un análisis más asiduo en las publicaciones y en las instrucciones de tipo político oficiales que, por ejemplo, en los informes agropecuarios, esa muestra sería representativa de las comunicaciones de masas disponibles, pero también tendenciosa respecto de las percepciones del círculo gobernante, precisamente a raíz de los aspectos de automuestreo implicados. Una de las probabilidades que debe sopesarse es la

del grado de visibilidad pública de los encargados de tomar las decisiones. Algunos individuos tienen un acceso más fácil a las comunicaciones públicas que otros, ya sea por el cargo que ocupan, por su jerarquía o por su personalidad. Otra probabilidad concierne a que las opiniones de alguien puedan o no reafirmarse en público. Las opiniones congruentes con las políticas establecidas, o que desempeñan funciones admitidas dentro del proceso político, tendrán más probabilidades de encontrar cabida en una publicación que las que se apartan de las finalidades generales.

Si se cuenta con una estimación de la probabilidad con la que un fenómeno que nos interesa está representado en el flujo de datos disponibles, una muestra no tendenciosa de estos datos debe asegurar que la frecuencia de dichos fenómenos en la muestra sea proporcional a la inversa de la probabilidad de que esos fenómenos estén disponibles para el muestreo. Si n_i es la frecuencia de los fenómenos i en la muestra, y p_i la probabilidad estimada de que el fenómeno i esté disponible para el analista, el criterio para medir la validez del muestreo es

$$n_i \text{ es proporcional a } p_i^{-1}$$

Cuando p_i tiene una distribución uniforme, y por lo tanto el automuestreo no está desviado, la validez del muestreo no hace sino demostrar que el muestreo de los datos disponibles fue aleatorio. En nuestro ejemplo, un muestreo válido de opiniones prestaría más atención a las personas encargadas de tomar decisiones que rara vez expresan sus opiniones en público, que a aquellos otros comunicadores de alta visibilidad que predominan en los medios de comunicación públicos. El examen llevado a cabo por Kops (1977, págs. 230-240) sobre los "diseños de muestreo aleatorio estratificado no proporcional" en el análisis de contenido revela que son los que más próximos están a un plan de muestreo satisfactorio de acuerdo con nuestro criterio sobre la validez del muestreo.

Con este ejemplo no se pretende sugerir que el análisis de contenido deba reproducir el resultado que ofrecerían normalmente otras técnicas. Por el contrario, el análisis de contenido debe tener en cuenta, por ejemplo, que algunos de los encargados de adoptar las decisiones tienen más influencia que otros, y aplicar esto para extraer una muestra que sea representativa de un universo de consecuencias posibles. O bien la naturaleza de las opiniones expre-

sadas podría dejarse de lado por completo, y en su lugar inferir la dinámica de las intenciones políticas que están detrás de una presentación pública preparada de antemano. En cada uno de estos casos, las muestras tal vez tengan que extraerse de manera diferente. En el análisis de contenido, *es importante demostrar la validez del muestreo, en especial cuando hay implicados procesos de automuestreo.* La definición del universo que nos interesa de modo que coincida con los datos disponibles, evita los problemas de muestreo, pero reduce tan seriamente la eficacia del análisis de contenido que convierte esta técnica en algo virtualmente inútil.

Validez correlacional

Un hecho muy reconocido por la bibliografía psicológica es que tanto *los resultados de los tests como los criterios de los tests son medidas,* y que ninguna de ellas debe confundirse con los fenómenos que podrían representar. Además, las correlaciones estadísticas evalúan *la intensidad de una asociación* entre dos o más variables, y por lo tanto indican el grado en que puede establecerse con certidumbre una relación (lineal) entre ellas. De ahí que las correlaciones no predigan sucesos o fenómenos, sino que indican (lo cual es muy importante en el desarrollo de instrumentos de medición) si una medida es *reemplazable* por otra. La validez correlacional tiene importancia por cuanto permite transferir la validez de una o más medidas establecidas a una medida nueva, que tenga ciertas ventajas prácticas. Así, los resultados de un análisis de contenido pueden hacerse más aceptables si se correlacionan con los tests psicológicos establecidos, las entrevistas, los experimentos, la observación directa, etcétera.

Campbell y Fiske (1959) fueron los primeros en desarrollar la idea de la validación mediante técnicas correlacionales hasta convertirla en una metodología cabal. Estos autores admiten que cualquier justificación de una nueva medida no sólo exige que ésta tenga una alta correlación con las medidas ya establecidas del rasgo que se intenta medir, sino también una correlación baja o nula con las medidas establecidas de los rasgos respecto de los cuales se intenta discriminar a aquél. El primer requisito se denomina validez *convergente*; el segundo, validez *discriminante*. Así, un resultado de investigación puede quedar invalidado o bien porque

la correlación con medidas independientes del mismo fenómeno sea demasiado baja, o bien porque la correlación con las medidas de los fenómenos respecto de los cuales se pretende diferenciar al primero sea demasiado alta, o por ambos motivos.

Para mostrar que una medida posee tanto validez convergente como validez discriminante, se requiere una matriz de rasgos y métodos múltiples de todas las correlaciones entre las medidas de una variedad de rasgos, cada una de las cuales debe haberse obtenido mediante métodos independientes. (Véase CAMPBELL y FISKE, 1959; ALWIN, 1974; KRIPPENDORFF, 1980a.)

En el campo del análisis de contenido, la validez correlacional reviste particular importancia cuando los fenómenos que interesan median entre la recepción y la producción de los mensajes. Esto se refiere, con mayor evidencia, a todos los conceptos mediadores del significado que están en la base de un gran número de diseños de investigación, y que se explicitan, entre otros trabajos, en el sistema de significado afectivo de Osgood. El primero en reconocerlo fue JANIS (1965), quien ya en 1943 había sugerido que la tarea del analista de contenido consiste en "estimar las significaciones atribuidas a los signos por un determinado público" (1965, pág. 61). Este autor concebía las "significaciones" [*significations*] como significados [*meanings*] representados internamente, que vienen de inmediato a la mente de un sujeto siempre que se enfrenta con un signo, aseveración verbal o símbolo capaz de afectar el comportamiento verbal o no verbal de los miembros del público. Janis puntualiza que las significaciones no son directamente observables. A raíz de los presuntos efectos de estas significaciones en la conducta de los receptores de un mensaje, para que un análisis de contenido sea válido sus resultados deben estar correlacionados, por lo menos, con algún aspecto de la conducta del público.

VALIDEZ PREDICTIVA

La predicción es un proceso mediante el cual se hace extensivo un conocimiento disponible a un campo desconocido. Los fenómenos predichos pueden haber existido en algún momento del pasado (como en el caso de los sucesos históricos o de las condiciones antecedentes de los mensajes disponibles), pueden ser contemporáneos de los datos que se analizan (como en el caso de las actitudes,

psicopatologías o personalidades de los sujetos entrevistados), o pueden anticiparse a algún hecho observable del futuro (como se supone en los tests de aptitud o en las extrapolaciones de tendencias). Aunque en las "Technical Recommendations" se distinguen los distintos tipos de validez de acuerdo con esta dimensión temporal, en el análisis de contenido esta distinción se subordina al hecho de que las inferencias específicas se efectúan respecto de fenómenos que están en el contexto de los datos disponibles.

No debe confundirse sustituibilidad con predecibilidad. Una medida es sustituible por otra si puede demostrarse que ambas ocupan la misma posición respecto de la red de correlaciones posibles. Y la validez correlacional evalúa justamente eso. En contraste con ello, una medida es predictiva de algún fenómeno si puede demostrarse que *concuerda* con lo que pretende predecir. En el caso ideal, las predicciones y los hechos guardan entre sí una correspondencia semántica biunívoca, es decir, toda predicción designa inequívocamente fenómenos en principio observables, ya sea directa o indirectamente, ya sea en el pasado, en el presente o en el futuro, y con independencia del modo en que llegan hasta el analista las pruebas respecto de tales fenómenos.

Uno de los aspectos de esta distinción radica en la diferencia entre acuerdo y correlación. Un reloj que atrasa mantiene una alta correlación con el tiempo, aunque a la vez tiene un desacuerdo sistemático. A menos que se conozca la desviación del reloj, no se podrá conocer la hora convencional. El famoso cómputo de cuerpos humanos [*body-count*] durante la guerra de Vietnam puede que mantuviera una alta correlación con la actividad militar, pero su valor numérico no mantenía ningún acuerdo con la cantidad de víctimas. Una alta correlación puede ser una condición necesaria, pero no suficiente, de la predecibilidad.

Un segundo aspecto de la distinción es el carácter de evidencia que se atribuye a la variable respecto de la cual se valida una medida. Al poner a prueba la validez correlacional, todas las medidas se consideran *índice de algún fenómeno subyacente* y, si existe sustituibilidad, son además cada una de ellas índice de la otra. Para poner a prueba la validez predictiva, se considera que la variable respecto de la cual se valida una medida es el *criterio final*, el "hecho", por así decirlo. Aunque también los "hechos" sólo son accesibles por vía de las observaciones que pueden ser descritas, y por ello no difieren fundamentalmente de las mediciones, lo que suministra el contenido de las predicciones es dicha

variable tal como es observada, y no lo que ella pueda indicar. Por ejemplo, tal vez a un gerente no le interese hasta qué punto las puntuaciones de los candidatos a ocupar un puesto en su empresa guardan correlación entre distintos tests de aptitud, pero sí cuál es el trabajo *efectivo* de esos candidatos una vez que inician su tarea. Y las medidas de la violencia en los programas televisivos, cualquiera que sea su correlación con otros índices, sólo obtienen validez predictiva cuando puede demostrarse su acuerdo con fenómenos observables (comportamiento del público, índice de delitos, temor de la población, expectativas de que se produzcan hechos violentos, etc.). Para probar la validez correlacional, puede servir como variable de criterio un índice; para probar la validez predictiva, el carácter fáctico del criterio es esencial. Cualitativamente, la validez predictiva queda asegurada ingresando cada uno de los sucesos de una serie de sucesos posibles en la siguiente tabla de cuádruple entrada, y midiendo luego el acuerdo que existe entre la predicción y la observación.

	sucesos predichos	sucesos excluidos por la predicción
sucesos que ocurrieron efectivamente	A	B
sucesos que no ocurrieron	C	D

Obviamente, los casilleros A y D obran en favor de la validez predictiva, mientras que los casilleros B y C obran en contra de ella. Ninguna columna debe estar vacía, pues de lo contrario no se podría poner en evidencia la discriminación, y tampoco ninguna fila debe estarlo, pues de lo contrario no podría demostrarse la convergencia.

Un ejemplo clásico, aunque no del todo concluyente, de esta forma de validación es el intento que llevó a cabo GEORGE (1959a) con respecto a evaluar las predicciones de la FCC norteamericana sobre la propaganda interna enemiga durante la segunda guerra mundial. Todas las inferencias se presentaron en forma de informes de los analistas de la propaganda, y podían equipararse

una a una con documentos a los que se tuvo acceso después de la guerra. Las inferencias para las cuales se contaba con pruebas validantes se juzgaron correctas, aproximadamente correctas o equivocadas. Los resultados mostraron que los trabajos del análisis habían tenido un considerable éxito predictivo. El intento de validación no fue del todo adecuado porque, al situar estas predicciones en esas tres categorías (o similares), no podían diferenciarse los casilleros B y C, mientras que el casillero D podía permanecer vacío.

En una modalidad más cuantitativa, la validación predictiva puede atenerse a los criterios de Campbell y Fiske, con una diferencia importante: que las entradas de la matriz de rasgos y métodos múltiples son *coeficientes de acuerdo*, y no correlaciones. (Sobre la naturaleza de estos coeficientes, véase el capítulo 12, "Fiabilidad".) De acuerdo con esto, puede decirse que un análisis de contenido tiene validez predictiva si se demuestra que sus inferencias exhiben tanto un alto acuerdo con los fenómenos (pasados, presentes o futuros) que pretenden predecir en el contexto de los datos, como un bajo acuerdo con aquellos otros fenómenos respecto de los cuales pretende establecer discriminaciones.

VALIDEZ DE LA CONSTRUCCION

Es probable que sea imposible efectuar la validación pragmática, ya que ello requiere experiencias anteriores de análisis del tipo de datos con que se cuenta y/o la disponibilidad de indicadores contemporáneos sobre los fenómenos en relación con los cuales se quieren formular inferencias. Y como la validez semántica y la validez del muestreo rigen sólo, en el caso típico, en las fases iniciales de un análisis de contenido, durante las cuales se mantiene la identidad de las unidades, la validez orientada a los datos no llega tan lejos como uno quisiera. Para los análisis de contenido que son idiosincrásicos respecto de un conjunto particular de datos, o que tienen precedentes en una situación dada, la validación de sus procedimientos como construcciones de la relación entre los datos y el contexto puede ser el único recurso disponible.

Por lo general, la labor de los historiadores es de esta índole. El enunciado "La historia nunca se repite" puede reflejar una posición filosófica o un fenómeno histórico, pero lo cierto es que también refleja la posición que asumen muchos analistas del conteni-

do. Cuando así sucede, quedan prácticamente descartados los procedimientos de validación estadística. DIBBLE (1963), que analizó los argumentos de los historiadores en favor o en contra de las inferencias extraídas a partir de documentos, llegó a la conclusión de que todos ellos implicaban *supuestos* acerca de las características psicológicas de los observadores, las reglas del sistema social que llevaba los registros, y las condiciones materiales y sociales que influyeron en la redacción de los documentos. Si, por un lado, los documentos históricos y las inferencias extraídas a partir de ellos se consideran idiosincrásicos y relevantes, los supuestos que conectan un texto con algún suceso tienen el carácter lógico de *generalizaciones sobre las pruebas documentales*.

Las inferencias efectuadas por LEITES y otros (1951) a partir de los discursos pronunciados en la conmemoración del nacimiento de Stalin (que ya hemos examinado) son otro ejemplo, particularmente transparente, de validación de la construcción en una situación esencialmente única. Las diferencias eran simplemente una derivación de la construcción, y la validez de ésta se estableció mediante cuidadosas argumentaciones para la generalidad de la relación, mostrando que su operacionalización captaba realmente dicha relación.

En un capítulo anterior mencionamos las cuatro fuentes de conocimiento a las que puede recurrir un analista de contenido para el desarrollo y justificación de sus construcciones analíticas:

- el éxito obtenido en el pasado con construcciones similares y/o en situaciones similares;
- las experiencias respecto del contexto de los datos disponibles;
- las teorías y modelos establecidos sobre las dependencias contextuales de los datos;
- la opinión de intérpretes y de especialistas representativos.

Estas fuentes son también las que probablemente haya que emplear para validar un determinado procedimiento analítico. La validación de una construcción abarca básicamente dos etapas:

1) Se procura *generalizar el conocimiento disponible* al contexto particular dentro del cual se efectúa el análisis de los datos, teniendo en cuenta: a) la relativa intensidad de las dependencias entre los datos y su contexto (por ejemplo, si las relaciones son deterministas o simplemente probabilísticas); b) la confianza en la validez de este conocimiento en lo que se refiere a la cantidad de

confirmaciones independientes, su correspondencia con la teoría establecida o la falta de ejemplos contrarios; y c) el grado en que este conocimiento se adecua al contexto particular, si es probable que las relaciones contextuales sean invariantes, y en qué forma el sistema que se examina difiere del conocido. Ya hemos estudiado todos estos puntos en el apartado sobre "Incertidumbre" del capítulo 9 ("Construcciones analíticas para la inferencia"), por lo que no es necesario que nos detengamos en otros pormenores.

2) Se procura *derivar lógicamente*, a partir de las generalizaciones válidas, las particulares *proposiciones subyacentes en el procedimiento analítico* utilizado. Para desarrollar nuevas construcciones analíticas, esto implica operacionalizar, o derivar, reglas justificables para la transformación de los datos a partir de una teoría determinada. En la validación de las construcciones ya existentes, significa poner a prueba si la construcción puede de hecho derivarse lógicamente de las generalizaciones, examinando, en especial: a) los errores de omisión, como las dependencias postuladas por las generalizaciones que la construcción no operacionaliza, y b) los errores de comisión, o las dependencias incorporadas a la construcción que no tienen fundamento en las generalizaciones.

Asimismo, es posible diseñar experimentos o crear otras situaciones a partir de los que se obtengan datos que correspondan directamente a la construcción. Un ejemplo, ya mencionado, es el trabajo realizado por OSGOOD (1959) para la obtención de pruebas validantes para el análisis de contingencia. Expuso a sus sujetos a diversos pares de nombres y comprobó que las asociaciones que establecían entre los nombres estaban influidas en grado significativo por su co-ocurrencia en el material estímulo, lo cual sugiere que una relación estadística de las co-ocurrencias en los medios de comunicación de masas puede llegar a predecir las asociaciones del público.

Volvamos a subrayar que la expresión "validez de la construcción", tal como nosotros la utilizamos, difiere un poco de la forma en que se emplea en las "Technical Recommendations". No obstante, esta diferencia no está relacionada con el espíritu de las presentes argumentaciones. En las aplicaciones que estas recomendaciones tienen en cuenta, "la validez de la construcción se evalúa... demostrando que ciertas construcciones explicativas dan cuenta hasta cierto punto del trabajo (de los individuos) en el test". Las "Technical Recommendations" conciben esta forma de validación

como un proceso doble: "En primer lugar, el investigador pregunta: a partir de la teoría (subyacente en el test) ¿qué predicciones haría con respecto a la variación de las puntuaciones de una persona a otra, o de una oportunidad a otra? En segundo lugar, recoge datos que confirmen dichas predicciones" (1954, pág. 14). Cuando un análisis de contenido es idiosincrásico, sólo la primera parte del proceso al que aluden esas recomendaciones puede completarse: la justificación del procedimiento y de las categorías de análisis a partir de la teoría válida. La mayoría de las situaciones que se presentan en el análisis de contenido dificultan el paso a la segunda etapa de recopilación de datos para confirmar las predicciones, etapa que hemos examinado al ocuparnos de la validez correlacional y predictiva.

14. Guía práctica

En este capítulo se esbozan los pasos que habitualmente siguen los analistas del contenido, desde la conceptualización del problema de la investigación hasta la confección del informe final. Un claro esquema de los procedimientos utilizados facilita la interpretación de los hallazgos y la reproducción del proceso que conduce a ellos.

Si en los capítulos precedentes hemos introducido conceptos analíticos y nos hemos extendido acerca de algunos problemas metodológicos, ahora intentaremos resumirlos pensando en su aplicación práctica.

Para empezar, nos encontramos con que el análisis de contenido abarca tres actividades lógicamente distintas:

- proyecto;
- ejecución;
- informe.

El orden lógico de estas actividades no implica, sin embargo, que la frecuencia temporal sea siempre ésta. En muchas ocasiones, el investigador resuelve adoptar un cierto esquema analítico a partir de una idea viable, y luego comprueba que no puede ejecutarlo de la forma en que lo había planeado. Ya sea porque tenga una

concepción errónea de sus datos, porque no haya podido encontrar la suficiente bibliografía como para justificar sus afirmaciones, o porque subestimara los recursos que necesitaba, tal vez se vea obligado a volver al trazado del proyecto hasta encontrar un diseño que sea factible y, *además*, conduzca a resultados que resulten significativos para él. También suele ocurrir que al redactar un informe sobre sus resultados, se le hacen evidentes ciertas pautas que exigen posteriores cálculos o una fundamentación de los controles de calidad, lo cual implica pasar de la etapa del informe a la etapa de la ejecución hasta formar un cuadro claro y cabal. Pese a estas idas y venidas, pese a estos ajustes relacionados con las restricciones existentes o a estas transacciones entre requisitos antagónicos, las prioridades lógicas se mantienen: el diseño al que finalmente se llegue explicitará lo que ha de ejecutar, y la ejecución producirá resultados de los que a la postre informará.

Como cabe afirmar de la mayoría de las investigaciones, los análisis de contenido rara vez acaban completándose. Aunque un buen análisis puede responder a algún interrogante, también es previsible que plantee otros nuevos que obliguen a revisar los procedimientos para las aplicaciones futuras, o que estimule nuevos estudios en lo que se refiere a los fundamentos de la formulación de inferencias, para no mencionar las nuevas hipótesis que puede sugerir el investigador sobre los fenómenos que le interesan. El comienzo y el final de un análisis de contenido sólo indican un período de tiempo arbitrario.

Proyecto

"Proyectar" o "diseñar" significa llevar a la práctica efectiva una idea y hacer operativa una manera de observar vicariamente la realidad. El desarrollo del proyecto de investigación es, desde el punto de vista intelectual, la más apasionante de las actividades que incluye el análisis de contenido; en ella, el investigador esclarece sus propios intereses y conocimientos, explora la bibliografía existente en búsqueda de intelecciones acerca de las condiciones circundantes de sus datos, juega con las ideas y las somete a pretests hasta que surgen planes en los que todas ellas confluyen en un procedimiento de investigación ejecutable (véase el capítulo 4, "La lógica del proyecto del análisis de contenido"). En la etapa del proyecto, la metodología se utiliza de manera creativa y con

fines prácticos. En lo que sigue, examinaremos ocho aspectos del proyecto de un análisis de contenido:

- Aplicación del marco de referencia conceptual al análisis de contenido.
- Búsqueda de los datos adecuados.
- Búsqueda del conocimiento contextual.
- Desarrollo de planes para la determinación de las unidades y el muestreo.
- Desarrollo de las instrucciones de codificación.
- Búsqueda de procedimientos justificados según el contexto.
- Establecimiento de las normas de calidad.
- Presupuestación y asignación de los recursos.

Aplicación del marco de referencia conceptual al análisis de contenido

Si toda investigación científica está motivada por el deseo de conocer o de entender mejor una porción del mundo real, un análisis de contenido debe interesarse por *dos especies de realidad*: la realidad de los datos y la realidad de lo que el investigador quiere conocer. En el análisis de contenido estas dos realidades no se superponen, no son coincidentes, de modo que el investigador tendrá que encontrar la manera de considerar los datos que está en condiciones de analizar como manifestaciones simbólicas o indicadores de los fenómenos que le interesan.

En el caso ideal, el analista comienza clarificándose a sí mismo lo que de verdad le interesa conocer y no puede observar de forma directa, y a continuación busca datos que le permitan extraer inferencias sobre la porción del mundo que le interesa. Esto significa, ante todo, determinar *el objetivo de las inferencias*.

Como ejemplo de estos objetivos, mencionemos las cuestiones concernientes a las percepciones del público, las psicopatologías, los servicios de inteligencia militar y la naturaleza de los acontecimientos históricos (véase el capítulo 2, "Fundamentos conceptuales"). Pero, en la práctica, muchos analistas parten del extremo opuesto, de los datos disponibles, y tratan luego de explorar su naturaleza simbólica para ver qué podrían inferir de ellos que les resultase interesante. A este último enfoque se le llama a veces, metafóricamente, "excursión de pesca". BERELSON (1952) y

muchos otros autores se han pronunciado contra él, no obstante lo cual los problemas metodológicos que plantea pueden sortearse en gran medida si el analista, tras su exploración inicial, distingue con claridad los datos de lo que desea inferir, y a continuación vuelve a examinarlos para ver si siguen siendo significativos, representativos y apropiados en el contexto de lo que a él le preocupa. En el análisis de contenido, la elección de los datos debe estar justificada por lo que el analista quiere saber; de ahí la prioridad lógica del objetivo.

Los resultados de una investigación deben ser, en principio, validables. Al trazar un objetivo para sus inferencias, el analista puede querer especificar, de manera inequívoca, la clase de *pruebas que demostrarían la validez* o la falta de validez de dichas inferencias. Esto puede ser obvio cuando se trata de inferir estados psicológicos o de identificar a autores desconocidos, pero puede ser difícil si se trata de inferir las actitudes de figuras históricas o de "predecir" la naturaleza de ciertos sucesos simbólicos que tuvieron lugar en una civilización desaparecida. No obstante, aun cuando defina una medida (por ejemplo, de la "prominencia simbólica") sobre la frecuencia con que aparecen ciertas palabras en un texto, el analista debe siempre aclarar qué es lo que pretende indicar con ella.

Búsqueda de los datos adecuados

Debe saberse previamente si los datos que se han utilizar para un análisis de contenido tienen alguna conexión con lo que el investigador quiere inferir. Esa conexión entre los datos y el objetivo de las inferencias del analista puede ser muy tenue, y de hecho rara vez es intensa y obvia desde el comienzo; pero no tendría ningún sentido reunir datos que pueden ser simbólicos o poseer significados en un contexto distinto, si esos significados nada tienen que ver con el interés del analista. Presumiblemente, el motivo para someter a análisis de contenido a los medios de comunicación de masas es que éstos, más allá del entretenimiento que puedan ofrecer, reflejan ordenamientos institucionales (socioeconómicos) de la sociedad, son elementos poderosos que moldean la opinión pública e incluso pueden guardar una conexión causal con diversas patologías sociales.

La búsqueda de datos adecuados puede ser posterior a cualquier vínculo, conocido o previsto, entre lo que se ha de inferir y

lo que es concebible que pueda observarse, muestrearse y analizarse. Los análisis de contenido tradicionales ponían el acento en las *referencias semánticas* y en las *expresiones* de las evaluaciones actitudinales, pero en los usos modernos los datos pueden concebirse como *correlatos* de los fenómenos de interés —por ejemplo, el desgaste de los libros de una biblioteca permite evaluar la popularidad de ciertos temas—, como *productos colaterales* de estos fenómenos (cuando se utiliza el contenido de un medio de comunicación de masas para inferir la conformación económica y las tendencias ideológicas de la industria de las comunicaciones), como *causas* (en el intento de inferir las percepciones del público o los planes de acción política) o como *elementos instrumentales* (cuando se entiende que los datos son manifestación de los valores o intereses particulares de quien los ha producido). WEBB y otros (1966) repasan dichas vinculaciones en función de las *huellas físicas* que dejan tras ellos los fenómenos de interés y de los *registros actuariales* que se mantienen por motivos distintos de los del interés que el analista tenga por ellos. Así, a fin de buscar datos adecuados, el analista debe examinar de qué manera los fenómenos que le interesan pueden influir en (o sufrir influencias de) lo que él puede observar, de qué manera se producen esos fenómenos y dónde pueden encontrarse huellas suyas o de sus respectivas representaciones lingüísticas. *Todo lo que esté relacionado con los fenómenos que le interesan al analista de contenido merece ser considerado como dato para el análisis.*

Búsqueda del conocimiento contextual

Las pruebas sobre la conexión empírica existente entre los datos y lo que se ha de inferir de ellos tiene una importancia obvia en cualquier análisis de contenido. Ya hemos aducido que una alta frecuencia de símbolos acerca del amor y el sexo no indica necesariamente que se esté en presencia de una sociedad promiscua; si lo que se tiene presente es una hipótesis de sustitución simbólica, más bien se inferirá lo opuesto, una sociedad represiva. Asimismo, la interpretación de estas frecuencias como índice de atención puede que no ofrezcan una corroboración lógica. Para que cualquier inferencia a partir de los datos esté justificada, es esencial contar con algún conocimiento riguroso, alguna prueba empírica sobre las conexiones entre éstos y lo que se ha de inferir de ellos. Dicho

conocimiento es lo que permite al investigador situar sus datos en un contexto adecuado, convertirlos en indicativos de fenómenos que estén más allá de ellos, y proporcionarle así un *puente lógico* para la formulación de inferencias. En esta búsqueda, el analista del contenido se convierte en un consumidor de saber.

Fuentes típicas de conocimiento contextual son: las *teorías y modelos* acerca del sistema que se investiga, las *experiencias con el contexto de los datos* que el analista puede querer operacionalizar, los *éxitos* que tuvieron en el pasado análisis de contenido con datos similares en situaciones similares, y los *intérpretes representativos*, como los especialistas o los informantes de las cualidades simbólicas de los datos (véase el capítulo 9, "Construcciones analíticas para la inferencia"). El investigador que diseña un análisis de contenido no debe sorprenderse de que la búsqueda de esta clase de conocimientos le absorba por lo general gran parte de su tiempo y esfuerzos. Las inferencias no se justifican por sí solas. Según el grado en que se hayan realizado buenas investigaciones en un cierto campo, habrá mayores o menores probabilidades de que existan muchos estudios que investiguen hipótesis capaces de justificar las construcciones analíticas para obtener las inferencias deseadas. Tal vez haya informes sobre análisis de contenido ya terminados que partieron de premisas similares. El éxito de dichos análisis puede favorecer la repetición de las mismas premisas. Tal vez haya expertos que conozcan o hayan asimilado las dependencias entre los datos y los fenómenos de interés. Ya hemos mencionado (véase el capítulo 9) algunas de las dificultades que plantea la aplicación de los conocimientos de esta índole.

Con suma frecuencia, la búsqueda de este conocimiento contextual obliga a realizar experimentos preparatorios, por ejemplo utilizando sujetos para probar si la teoría "funciona" en contextos semejantes a los del análisis. Ejemplos de ellos son los experimentos de validación del análisis de contenido llevados a cabo por OSGOOD (1959) y los que efectuaran STONE y HUNT (1963) sobre las notas escritas por suicidas. En este caso, las correlaciones entre las frecuencias de símbolos clave y la atención o importancia asignada a sus referentes por los autores o los lectores fueron sorprendentemente bajas. En vista de lo poco que se conoce acerca de las relaciones entre el uso del lenguaje y las expresiones empleadas en los medios de comunicación, por un lado, y las condiciones antecedentes, correlatos conductuales y consecuencias,

por el otro, es necesario explorar más sistemáticamente esas relaciones.

Desarrollo de planes para la determinación de las unidades y el muestreo

Dado un universo de datos posibles, o al menos un esbozo, el investigador debe encontrar el modo de obtener todos esos datos o una muestra. Si dispone de plena libertad para incluir o excluir un dato particular en su análisis, puede atenerse a algunos de los procedimientos de muestreo corrientes.

No obstante, en la mayoría de los análisis de contenido este control sobre el proceso de recopilación de datos es limitado. Normalmente, hay buenas razones por las cuales algunos datos son más accesibles al analista que otros. La supervivencia de documentos históricos es selectiva, y también lo son las declaraciones de los pacientes ante sus psiquiatras y lo que se imprime como noticia en un momento dado: todo ello se selecciona a partir de una amplia gama de posibilidades. Una selectividad de esta índole nos está indicando procesos de muestreo inherentes a la fuente. Ya sea que dichos procesos sean inconscientes (como quizás, en las entrevistas psiquiátricas) o estén regidos por decisiones racionales (como tal vez, en la programación televisiva y en la publicidad), lo cierto es que el analista sólo puede escoger entre lo que ha sobrevivido a los procesos de automuestreo. Y al hacer esta elección, tal vez deba rectificar la tendencia del automuestreo inherente a los datos disponibles.

El conocimiento necesario para el diseño de planes destinados a la determinación de las unidades y al muestreo abarca:

- la índole de las unidades de muestreo que pueden ser portadoras de información de manera individual (véase el capítulo 5, "Determinación de las unidades");
- la localización espacial o temporal de unidades de distinta clase (véase el capítulo 6, "Muestreo");
- el tipo de distribución de la información en el universo;
- las características en materia de automuestreo de distintas clases de unidades de información (véanse los capítulos 6, "Muestreo", y 13, "Validez").

Es preciso recurrir a procesos aleatorios para la elección final en cuanto a la inclusión de datos en la muestra, pero sólo después de haber agotado todo el conocimiento con que se cuente sobre el universo de datos posibles.

En la práctica, todo plan de muestreo debe afrontar gran cantidad de irregularidades. Por ejemplo, se solicitan números atrasados de periódicos a su editor y ocurre que ya no se puede contar con ellos; se seleccionan para su análisis ciertos fragmentos de entrevistas grabadas y ocurre que contienen demasiado ruido; se pretende hacer un estudio de las características de la personalidad a partir de una muestra de espectáculos televisivos y ocurre que en dichos espectáculos no aparece ningún actor. En todas estas circunstancias, el investigador deberá establecer: 1) si la deficiencia es totalmente accidental, en cuyo caso puede continuar con el proceso de elección al azar hasta que la muestra tenga el tamaño apropiado; ó 2) si la deficiencia es el resultado de tendencias del automuestreo, en cuyo caso tendrá que empeñarse por obtener la unidad que falta y, si no lo logra, indagar por qué ha quedado fuera esta unidad y por cuál otra puede reemplazarse si se quiere conservar la representatividad de la muestra; ó 3) si la deficiencia es el resultado de una conceptualización inadecuada, ya sea del universo de datos posibles o de su estratificación, agrupamiento u organización. Ello exige revisar el plan de muestreo, o quizá lo que constituye el universo de datos posibles para el estudio.

El plan de muestreo tiene que ser minucioso y explícito para convertirse en un procedimiento reproducible y dar lugar a muestras similares a partir de un mismo universo.

Desarrollo de las instrucciones de codificación

Muchos investigadores parecen apreciar en alto grado sus propias categorías de análisis, y entonces formulan y someten a pretests las nuevas instrucciones para los codificadores hasta que son lo bastante fiables, y luego las aplican a datos que jamás han sido analizados antes.

De esta manera, se emprenden gran cantidad de análisis de contenido cuyos resultados no son comparables entre sí ni tampoco contribuyen acumulativamente al desarrollo teórico. Aunque la originalidad es siempre bienvenida, el investigador que se base en las conceptualizaciones ya existentes tiene más probabilidades de efectuar esa contribución. Un punto de partida obvio es examinar

la bibliografía disponible acerca del modo en que los datos están relacionados con su contexto. No todas las características descriptibles de un mensaje se vinculan con los fenómenos que interesan, y por ello tiene poco sentido ampliar un análisis de contenido más allá de su finalidad. Habitualmente, existen teorías bien formuladas o, por lo menos, hipótesis capaces de suministrar conceptos y categorías en cuyos términos es posible describir nuevos datos. Esto le permitirá al investigador confiar a partir de ese momento en construcciones bien elaboradas. También deberá confiar, en lo posible, en las categorías utilizadas en otros estudios. Por ejemplo, si un análisis de contenido requiere categorías ocupacionales, puede basarse en las listas oficiales de la Federal Trade Commission de los Estados Unidos u otros organismos análogos, o en estudios sociológicos sobre el nivel ocupacional y el prestigio. Esto posibilita que los datos sean comparables. La dificultad corriente consiste en convertir en operativos conceptos y categorías teóricos procedentes de un campo distinto, de modo que se amolden a un lenguaje de datos adecuado, con sus variables y sus categorías mutuamente excluyentes (véase el capítulo 8, "Lenguajes de datos") por un lado, y por otro a las concepciones espontáneas de la fuente de los datos, y sigan siendo comprensibles para los codificadores humanos o las aplicaciones mediante ordenador.

Una segunda fuente de ideas para la codificación de las instrucciones son *los análisis de contenido publicados que tengan finalidades similares.* En los capítulos precedentes ya se han ilustrado diversos sistemas de categorías (en especial en el capítulo 7, "Registro", y en el capítulo 8, "Lenguajes de datos"); otros ejemplos pueden encontrarse en BERELSON (1952) y particularmente en HOLSTI (1969). Hay estudios que se basan en unas pocas variables, mientras que otros definen muchísimas; algunos sólo requieren una página de instrucciones, mientras otros necesitan un libro entero (véase DOLLARD y AULD, 1959). Pero no hay necesidad alguna de inventar un nuevo esquema si no se ha demostrado la inadecuación de los existentes. Y aun cuando la validez de estos instrumentos no esté del todo establecida, el uso de instrucciones de registro ya existentes o de esas mismas instrucciones una vez sometidas a leves adaptaciones, ofrece dos netas ventajas sobre el uso de nuevos esquemas: en primer lugar, proporciona la posibilidad de comparar los resultados en distintas situaciones, a través de lo cual pueden surgir normas o expectativas (véase el capítulo 3, "El uso de la inferencia y sus clases") y, en segundo

lugar, abrevia los esfuerzos destinados a convertir en fiables esas instrucciones.

Las instrucciones de codificación o de registro, para ser adecuadas, deben:

- prescribir las características que han de reunir los observadores (codificadores, jueces) empleados en el proceso de registro;
- establecer la capacitación a que deben someterse esos observadores a fin de prepararse para la tarea;
- definir las unidades de registro, incluyendo los procedimientos para su identificación;
- delinear la sintaxis y la semántica del lenguaje de datos (variables, categorías), incluyendo en caso necesario un esbozo de los procedimientos cognitivos utilizados para situar los datos dentro de las categorías;
- describir de qué manera se han de utilizar y administrar las planillas de datos.

Desde el momento en que las instrucciones finales de codificación sirven de base para reproducir el análisis en otra oportunidad o para aplicarlo a datos similares, dichas instrucciones deben ser explícitas y detalladas, y si no forman parte del informe final de la investigación, al menos deben estar disponibles en caso de solicitárselas.

Búsqueda de procedimientos justificados según el contexto

En el capítulo 2, sobre los "Fundamentos conceptuales" de un análisis de contenido, desarrollamos el postulado de que su objetivo está más allá de la sintaxis y forma de los datos, cuya naturaleza simbólica es mediadora respecto de algo perteneciente al contexto de dichos datos, al tiempo que lo señala o lo indica. Por consiguiente, los procedimientos analíticos han de ser sensibles a este contexto y representarlo, o justificarse en función suya. El análisis de contingencia (véase el capítulo 10, "Técnicas analíticas") es un ejemplo de un procedimiento de cómputo que operacionaliza ciertas proposiciones psicológicas y, por lo tanto, es sensible al contexto en que dichas proposiciones resultan verdaderas. También el análisis de la aseveración evaluativa deriva de ciertas teorías sobre la cognición. En otros capítulos (particularmente en el capítulo 9, "Construcciones analíticas para la inferencia" y en el

13, "Validez") nos extendemos sobre las dificultades de la operacionalización y sobre los criterios que deben satisfacer las construcciones analíticas.

Debe subrayarse que *todo procedimiento analítico, por su propia naturaleza, implica ciertos supuestos acerca del contexto de los datos que deben defenderse a partir de lo que se conoce sobre este contexto.* Cuando se examinan los supuestos implícitos en los procedimientos utilizados por los analistas de contenido (por lo general sin una justificación explícita), sorprende a menudo comprobar hasta qué punto la concepción enunciada por el analista sobre el entorno simbólico se aparta de la concepción incorporada a la técnica analítica que ha escogido.

Establecimiento de las normas de calidad

En gran parte, el análisis de contenido se puede caracterizar como un juego en que el analista trata de conjeturar qué es lo que le oculta su contrincante; y para obtener un mejor resultado que si se dejase guiar por el azar, conviene que confíe en cualquier información que dicho contrincante pueda eventualmente revelarle sobre él mismo. Es probable que en ese proceso se produzcan errores, y éstos pueden ser costosos. En las situaciones prácticas, en las que puede verse implicado el bienestar de las personas (véase el apartado sobre la validación), la certidumbre con que se formulan las inferencias tiene que ser mayor que en una obra académica, donde sólo está en juego la fama del investigador. En general, cuanto mayor es el costo de los errores, más costoso resulta el análisis de contenido, ya sea porque durante el análisis se requiere contar con mayor cantidad de datos, o ya por la necesidad de técnicas más perfeccionadas, entre ellas estudios preparatorios de las construcciones analíticas implicadas. Así pues, aunque siempre es conveniente que existan altas normas de calidad, éstas son también costosas, y un analista de contenido puede que no desee aceptar un proyecto cuyas normas sobrepasan lo que él está en condiciones de hacer, tanto en lo referente a los recursos de que dispone como al costo de los errores en los resultados.

Es importante, sin embargo, establecer estas normas *antes* de evaluar un análisis. En el análisis de contenido las normas cualitativas se refieren principalmente a la fiabilidad y la validez. El investigador las descubre yendo de adelante hacia atrás, desde la validez (exactitud, certidumbre o especificidad) que se exige de

los resultados, hasta las condiciones que cada elemento del análisis debe cumplir para asegurar la obtención de resultados válidos. Con este objeto, debe conocer el efecto neto de las normas individuales sobre el resultado.

Dado que un alto grado de fiabilidad es condición previa de un alto grado de validez, pero no la asegura, las normas sobre validez son evidentemente más imperiosas, y por ello preferibles a las normas sobre fiabilidad. No obstante, medir la fiabilidad es importante por dos motivos: en primer lugar, cuando se carece de pruebas validantes, el investigador no tiene otra opción que utilizar todo lo que pueda asegurarle que sus resultados son válidos, y esto es precisamente lo que hace la fiabilidad; en segundo lugar, dado que una baja fiabilidad significa que no es posible dar los datos por seguros, de nada sirve proceder a una prueba de validez, por lo general más costosa, si la fiabilidad es baja. De esta manera, la fiabilidad indica si los trabajos de validación pueden suministrar alguna seguridad adicional.

Presupuestación y asignación de los recursos

Un análisis de contenido se compone de múltiples y muy distintas clases de actividades. Algunas de éstas puede realizarlas una sola persona, mientras que otras exigen la colaboración de varias, como cuando se desea verificar la fiabilidad de las instrucciones de registro. Algunas llevan mucho tiempo, y otras apenas unas fracciones de segundo, como el procesamiento de datos por ordenador. Algunas deben ser ejecutadas en serie, mientras que otras pueden disponerse "en paralelo". Algunas exigen instalaciones y equipos escasos, y de esa manera demoran la ejecución de otras. Algunas sólo puede llevarlas a cabo el propio investigador, mientras que otras las puede delegar en ayudantes bien adiestrados. A menos que el análisis de contenido sea de poca envergadura y de carácter exploratorio, el investigador debe formarse una idea clara acerca de la organización del trabajo necesaria para llevar a cabo el análisis: cuándo y dónde se necesitan los recursos, por cuánto tiempo y a qué costo.

Un método útil consiste en establecer un cursograma. A partir de un cuadro general, se detallan las actividades implicadas en la manipulación de los datos, desde el muestreo hasta la redacción del informe. Esto obliga al investigador a prever los elementos que precisará en cada etapa.

Otras herramientas útiles para la investigación operativa son el PERT (Program Evaluation and Review Technique [Técnica de Evaluación y Revisión de Programas]) y el CPM (Critical Path Method [Método del Camino Crítico]) (ELMAGHRABY, 1970). Con estos métodos se convierte todo un proceso de investigación en una red de flechas, cada una de las cuales representa una actividad (como la perforación de tarjetas o la clasificación) y conecta entre sí dos nodos o estados de la investigación (por ejemplo, planillas de datos completadas y datos perforados, análisis completado, etc.). Para cada flecha se consignan el tiempo o los recursos necesarios para ejecutar el proceso. Estos métodos permiten al investigador planear las diversas actividades y averiguar cuál es la forma más eficiente de concluir el estudio, estableciendo así una organización del trabajo y pudiendo evaluar los requisitos en materia de recursos humanos para que el plan sea viable.

Por supuesto, existen numerosas técnicas de planificación que pueden consultarse. Cualquiera de ellas que se utilice, debe dar al analista de contenido, y quizá también a los organismos que financian la investigación y a sus clientes, las siguientes seguridades:

- Existe por lo menos una forma práctica de proceder (más adelante pueden surgir desviaciones respecto de ella que resulten más fructíferas, pero el diseño debe indicar por lo menos una). El investigador necesita conocer en cada punto exactamente dónde se encuentra y a dónde se dirigirá a continuación.
- La investigación puede llevarse a cabo de manera organizativa, en lo referente al personal disponible, sus conocimientos teóricos y prácticos, las instalaciones existentes, etc. El investigador necesita conocer dónde y cuándo las personas que participan realizan las diversas actividades, quién de ellas es la responsable o de dónde proceden los servicios.
- La investigación puede llevarse a cabo dentro de un período de tiempo razonable. Siempre que se plantee la cuestión, el investigador necesita conocer cuánto tiempo le llevará, y si está dentro de lo planeado o se excede.
- Hay suficientes recursos disponibles como para completar el análisis de contenido, y pueden subdividirse en costos de los sueldos y salarios, instalaciones y servicios, por hora, mes o año, o en función de presupuestos individuales. Sea lo que fuere lo que parezca más conveniente, el investigador debe poder evaluar, en cualquier momento del proceso, si se está desenvolviendo bien y en qué medida, cuánto dinero ha gastado y cuánto le queda aún.

EJECUCION

En un caso ideal, la ejecución de un análisis de contenido bien proyectado sería un asunto de rutina; si todos los problemas de corte intelectual se resolvieran durante la fase del proyecto, la ejecución podría delegarse a una entidad que se ocupase de investigaciones. Pero en la realidad no hay duda de que surgirán problemas imprevistos incluso en el plan de investigación mejor concebido. Muchos de estos problemas son producto de las insuficiencias en los datos disponibles y del costo imprevisto de los recursos humanos y las instalaciones; pero los más graves provienen de la incapacidad para satisfacer las normas preestablecidas de fiabilidad y validez. Cuando la solución a estos problemas no puede establecerse de antemano, en lugar de interrumpir por completo el análisis de contenido hay que volver atrás y modificar el proyecto, teniendo presente el objetivo global de la investigación.

El orden en que se ejecuta por lo general un análisis de contenido ya se ha esbozado en el capítulo 4, sobre "La lógica del proyecto del análisis de contenido". Ya sea que el estudio se efectúe a mano o por ordenador, o ya sea que su énfasis recaiga en la descripción o en las inferencias elaboradas, debe incluir al menos uno, y en lo posible más, de los siguientes aspectos:

- Muestreo mediante unidades hasta que la muestra pueda considerarse suficientemente representativa del universo.
- Identificación y descripción de las unidades de registro, que deben ser reproducibles y satisfacer los criterios de validez semántica donde éstos se apliquen.
- Reducción y transformación de los datos, dándoles la forma que exige el análisis a la vez que se retiene toda la información relevante.
- Aplicación de procedimientos analíticos sensibles al contexto (construcciones analíticas) para obtener las inferencias.
- Análisis, identificación de las pautas dentro de las inferencias, verificación de las hipótesis concernientes a las relaciones entre éstas y los resultados obtenidos mediante los métodos, y validación pragmática de los hallazgos.

Cuando se tiene la impresión de que no se va a poder reunir una muestra representativa, ya sea a raíz del alto costo que implica el gran tamaño de la muestra o, más habitualmente, porque hay ciertos ítems que simplemente son inaccesibles, poco puede

hacerse más allá de limitar las inferencias que se pretendían extraer de los datos a las que puedan obtenerse con algún grado de validez. Los individuos y las sociedades se parecen en su capacidad de eliminar la información que podría ser valiosa para un análisis de contenido, y aquello a lo que no se tiene acceso no puede examinarse. Un remedio consiste en redefinir el universo; otro en reducir el nivel del análisis de contenido al de un examen preliminar, hasta que puedan incluirse los datos relevantes.

Si el registro de los datos no puede llevarse a cabo de forma fiable, la primera tarea será localizar las fuentes de la falta de fiabilidad (véase el capítulo 12); ésta posiblemente tenga su origen en el hecho de que algunos codificadores no hayan sido bien adiestrados, o quizá sólo se produzca entre algunas categorías mal definidas, mientras que las restantes pueden ser claras e inequívocas; en fin, la falta de fiabilidad puede limitarse a unas pocas variables, mientras que el resto resulta aceptable tal como es.

Una vez localizada la fuente de la falta de fiabilidad, hay, en esencia, dos alternativas. En primer lugar, *omitir del análisis los aspectos poco fiables*. Por ejemplo, los datos registrados, categorías o variables que no sean fiables se eliminan de todo análisis posterior. Los datos que sobreviven a este proceso de selección incluyen entonces menos información que antes, y el investigador ha de resolver si las omisiones efectuadas no han introducido cierta tendenciosidad en la información retenida y si ésta es suficiente como para que sea justificable continuar con el análisis.

Pero más habitual es la segunda alternativa: *recodificar los datos mediante otros instrumentos* que dejen de lado las causas de la falta inicial de fiabilidad. Esto significa modificar las instrucciones de registro proporcionadas a los codificadores, volver a adiestrar a estos últimos, y administrar nuevas pruebas de fiabilidad hasta satisfacer todos los criterios. De hecho, no es común que surjan instrucciones de registro fiables sólo después de haber fracasado en dos o tres intentos previos. Un buen proyecto de investigación debe prever el perfeccionamiento continuo de las instrucciones de registro. Al capacitar a los codificadores debe intentarse codificar, en las instrucciones de registro, el contenido del proceso de capacitación; y al verificar la fiabilidad, el investigador debe asegurarse de que los codificadores resulten independientes entre sí, y con respecto a las tareas de codificación asignadas previamente (véase el capítulo 7, "Registro", y el capítulo 12, "Fiabilidad").

Cuando el texto se procesa directamente mediante ordenador, en cuyo caso las unidades (palabras, frases, oraciones) se identifican y categorizan por vía de algoritmos, sin intervención alguna de juicios humanos, la fiabilidad deja de ser un problema, pero puede cuestionarse la validez semántica del proceso. A fin de alcanzar niveles adecuados de validez semántica, es común también recorrer varios ciclos hasta que los procesos (diccionarios y rutinas correspondientes en el caso de la codificación mediante ordenador, y categorías de la instrucción de registro en el caso de la codificación manual) se ajusten a las normas deseadas.

La reducción de los datos mediante su selección de acuerdo con algún criterio, y su transformación, como cuando se copian trasladándolos de un medio de transmisión a otro, así como la tabulación y perforación con códigos, son tareas "burocráticas" que suelen pasarse por alto, pero que también están sujetas a errores humanos cuya magnitud a menudo excede a la del registro. Es preciso controlar todas estas actividades para asegurarse de que los datos sean exactos y de que no se haya omitido en el análisis nada que resulte relevante. Limpiar lo que suele denominarse datos "sucios" es una tarea que exige tiempo, y en la cual se cometen algunos errores típicos que deben controlarse, como los siguientes:

1) *Registros ausentes.* A veces se omiten imágenes de tarjetas íntegras del conjunto, lo cual puede originar una lectura errónea de casi todos los datos.
2) *Registros que no aparecen en la secuencia adecuada.*
3) *Valores ilegítimos* son las etiquetas de categorías que no tienen significado alguno en el lenguaje de los datos, tal como se definió en las instrucciones de codificación. Por ejemplo, un "5" carece de legalidad si la escala va de "+3" hasta "-3"; una letra es ilegítima si lo previsto son valores numéricos; así como también la ausencia de un valor cuando en todos los casilleros debe aparecer uno.
4) *Incongruencias,* que se presentan cuando una unidad de análisis es descrita en términos que se excluyen entre sí; por ejemplo, "mujer soltera embarazada", "niño de 999 años"...
5) *Improbabilidades,* que se presentan cuando los valores son legítimos dentro del contexto del instrumento de medida, pero están fuera de la gama de las expectativas; por ejemplo, un espectáculo televisivo violento pero en el que no figure ningún personaje humano (si la violencia se definió en función de los actos intencionales de destrucción); o bien un "profesor universitario adolescente".

Para la detección de los errores de tipo 1 y 2 es importante asignar números de secuencia a los casos y números de identificación a las tarjetas. Es recomendable que los investigadores procuren establecer una organización de los datos que sea fácilmente controlable. Los registros ausentes pueden conducir a una lectura de los datos fuera de secuencia, dando resultados absolutamente ininterpretables. Los errores del tipo 3 y 4 son fáciles de identificar mediante programas especiales, que imprimen todos los valores que están fuera del intervalo así como las co-ocurrencias ilegítimas. Aunque al investigador no le será difícil corregir estos errores, dicho control no lo pone a salvo de otros que se encuentran dentro del intervalo de valores posibles o dentro de lo que es lógicamente posible. Una alta frecuencia de errores del tipo 3 y 4 sugiere que los datos son más "sucios" de lo que es posible controlar sin remontarse a su fuente. Las improbabilidades (tipo 5) no son errores en sí mismas, porque puede ser posible, por ejemplo, producir un espectáculo de televisión violento sin incluir en la grabación a actores humanos. Estas improbabilidades sólo señalan la necesidad de someter a un examen posterior el material que ha servido de fuente.

INFORME

El informe de una investigación es una descripción autorizada de lo realizado, de los motivos por los cuales se hizo, de lo que se logró con ello y de su contribución al conocimiento existente. En medios sumamente institucionalizados, donde la motivación es bien conocida y los procedimientos están codificados, el informe puede limitarse a aquello que se ha apartado de los procedimientos corrientes. Quizás a la persona que se basará en él para tomar decisiones sólo le interesen los resultados, y confíe en la buena reputación del investigador y en que ha realizado bien su tarea. Sin embargo, en ningún trabajo científico se aceptan los hallazgos con independencia del procedimiento mediante el cual se obtuvieron, y este procedimiento debe poder reproducirse en otros lugares y oportunidades. Un informe de investigación necesita incluir específicamente alguno de los siguientes aspectos o todos ellos:

1) Una enunciación del *problema general* que aborda la investigación. Debe procurar transmitir al lector la importancia teórica o práctica

de resolver dicho problema, en relación con lo que actualmente se conoce y/o con los beneficios que podría suponer resolverlo.

2) Una exposición de *los antecedentes del problema*, que incluya una reseña de la bibliografía, en caso de disponer de ella, sobre la forma en que ha sido abordado el problema en el pasado, lo que han demostrado otros trabajos de investigación semejantes y los motivos por los cuales se supone que un análisis de contenido puede suministrar hallazgos interesantes. (Véase el apartado "Distinciones" en el capítulo 2, "Fundamentos conceptuales".)
3) Una enunciación de *los objetivos específicos del análisis de contenido*, que gobiernan la elección de los datos, métodos y diseños, y el vínculo entre esos objetivos y el problema general antes enunciado. Específicamente, debe resumirse el sistema que se ha de examinar, los datos, su contexto y el objetivo de las inferencias, todo ello según lo aprecia el análisis de contenido (véase el apartado "Marco de referencia conceptual" en el capítulo "Fundamentos conceptuales"), incluidas tal vez las restricciones empíricas bajo las cuales se emprende la investigación. Aquí se debe establecer, o al menos esbozar, qué clase de prueba será aceptada para invalidar las inferencias que se efectúen (véase el capítulo 13, "Validez").
4) Una justificación de *los datos, métodos y diseño elegidos*. Esto exige cotejar las premisas incorporadas a las construcciones analíticas y a las técnicas, con lo que se sabe sobre el contexto de los datos (ya sea en forma de teorías, de éxitos obtenidos en el pasado o de experiencias prácticas). (Véase el capítulo 9, "Construcciones analíticas para la inferencia".)
5) Una descripción de los *procedimientos efectivamente adoptados* (véase el capítulo 4, "La lógica del proyecto del análisis de contenido"), de modo que la investigación pueda ser reproducida por otros. Esto incluye una descripción de los *esquemas de determinación de unidades* (véase el capítulo 5, "Determinación de las unidades"), el *plan de muestreo* (véase el capítulo 6, "Muestreo") las instrucciones de registro, los planes de capacitación, una muestra de las planillas de datos (véase el capítulo 7, "Registro", y el 8, "Lenguajes de datos"), las técnicas utilizadas para el manejo y el análisis de los datos (véase el capítulo 10, "Técnicas analíticas") y, muy particularmente, *los resultados de las pruebas de fiabilidad* (véase el capítulo 12, "Fiabilidad") y de todos los esfuerzos realizados para validar algunas de las etapas del procedimiento o éste en su totalidad (véase el capítulo 13, "Validez").
6) Una presentación de los hallazgos y de su significación estadística (la bondad del ajuste en los casos en que corresponda, las puntuaciones de fiabilidad asociadas con cada uno y una evaluación de su probable validez). En aras del progreso del saber, es importante

asimismo dar cuenta de los métodos de análisis que fracasaron o que no produjeron hallazgos interesantes. La aparente falta de estructura puede ser tan importante como su presencia, y no debe suprimirse una en favor de la otra.

7) Una *evaluación autocrítica* de los procedimientos adoptados y de los resultados obtenidos, en relación con los objetivos enunciados (¿se ha alcanzado la meta perseguida?), con los costos y beneficios de sus resultados (¿ha valido la pena el esfuerzo?) o en contraste con lo que se conoce por otros métodos (¿ha resultado adecuado efectuar un análisis de contenido?). Una vez completado un proyecto, habitualmente el analista ha acumulado numerosas experiencias, intelecciones y advertencias que querrá compartir con sus colegas. La mayoría de las investigaciones, si por un lado suministran algunas respuestas a las preguntas iniciales, también plantean nuevos problemas para indagaciones futuras que el investigador no debe guardarse para sí.

No se pretende que esta lista sea exhaustiva. Es bien sabido que los analistas de contenido inventan ingeniosos artificios con el fin de obtener inferencias aparentemente válidas; así pues, habrá analistas que quizá quieran informar sobre actividades no mencionadas aquí. Tampoco se pretende prescribir con esta lista una forma canónica para los informes acerca de los análisis de contenido. Algunas descripciones pueden relegarse a un apéndice (listas de datos, hallazgos de los pretests, esquemas de categorías complejos), o reemplazarse por la referencia a otras publicaciones (descripción de las técnicas analíticas, tests corrientes, justificaciones dadas por otros autores), y quizás algunas se pongan a disposición de todos los que manifiesten un especial interés por ellas (datos, instrucciones de codificación escritas que pueden llegar a ser muy voluminosas, tiras impresas por el ordenador). Un analista de contenido está obligado a dar a conocer todo lo que ha hecho, si no en una publicación, al menos para todos los individuos que deseen utilizar sus hallazgos, reproducir el análisis o perfeccionar sus técnicas.

Los resultados de un análisis de contenido deben ser lo suficientemente sólidos como para correr el riesgo de que sus procedimientos sean objeto de una crítica bien informada y basada en pruebas empíricas, demostrando que evitar esa amenaza constituye una de las finalidades importantes de su metodología. En todo lo anterior se ha procurado exponer la lógica que preside la demostración de la validez de los hallazgos de un análisis de contenido.

Referencias bibliográficas

ABELSON, R.P. (1963): "Computer simulation of hot cognition", págs. 277-298, en S.S. Tomkins y S. Messick (comps.), *Computer Simulation of Personality*, Nueva York, John Wiley.

—— (1968): "Simulation of social behavior", págs. 274-356, en G. Lindzey y E. Aronson (comps.) *The Handbook of Social Psychology*, Reading, MA., Addison-Wesley.

ADORNO, T.W. (1960): "Television and the patterns of mass culture", págs. 474-488, en B. Rosenberg y D.M. White (comps.), *Mass Culture*, Nueva York, Free Press.

ALBIG, W. (1938): "The content of radio programs–1925-1935", *Social Forces* 16: 338-349.

ALBRECHT, M.C. (1956): "Does literature reflect common values?", *American Sociological Review* 21: 722-729.

ALLEN, L.E. (1963): "Automation: substitute and supplement in legal practice", *American Behavioral Scientist* 7: 39-44.

ALLPORT, G.W. (1942): *The Use of Personal Documents in Psychological Science*, Nueva York, Social Science Research Council.

—— (comp.) (1965): *Letters from Jenny*, Nueva York, Harcourt Brace Jovanovich.

—— y J.M. FADEN (1940): "The psychology of newspapers: five tentative laws", *Public Opinion Quarterly* 4: 687-703.

ALWIN, D.F. (1974): "Approaches to the interpretation of relationships in the multitrait-multimethod matrix", págs. 79-105, en H.L. Costner (comp.), *Sociological Methodology* 1973-1974, San Francisco, Jossey-Bass.

American Psychological Association (1954): "Technical recommendations for psychological tests and diagnostic techniques", *Psychological Bulletin Supplement* 51, 2: 200-254.

ARMSTRONG, R.P. (1959): "Content analysis in folkloristics", págs. 151-170, en I. de Sola Pool (comp.), *Trends in Content Analysis*, Urbana, University of Illinois Press.

ARNHEIM, R. y M.C. BAYNE (1941): "Foreign language broadcasts over local American stations", págs. 3-64, en P.F. Lazarsfeld y F.N. Stanton (comps.), *Radio Research 1941*, Nueva York, Duell, Sloan & Pearce.

ASH, P. (1948): "The periodical press and the Taft-Hartley Act", *Public Opinion Quarterly* 12: 266-271.

ASHEIM, L. (1950): "From book to film", págs. 299-306, en B. Berelson y M. Janowitz (comps.), *Reader in Public Opinion and Communication*, Nueva York, Free Press.

BALDWIN, A.L. (1942): "Personality structure analysis: a statistical method for investigating the single personality", *Journal of Abnormal and Social Psychology* 37: 163-183.

BALES, R.F. (1950): *Interaction Process Analysis*, Reading, MA, Addison-Wesley.

BARCUS, F.E. (1959): "Communications content: analysis of the research 1900-1958: a content analysis of content analysis", Tesis, University of Illinois.

BARTON, A.H. (1968): "Bringing society back in: survey research and macromethodology", *American Behavioral Scientist*, 12, 2:1-9.

BATESON, G. (1972): *Steps to an Ecology of Mind*, Nueva York, Ballantine.

BECKER, H.P. (1930): "Distribution of space in the *American Journal of Sociology*, 1895-1927", *American Journal of Sociology* 36: 461-466.

—— (1932): "Space apportioned forty-eight topics in the *American Journal of Sociology, 1895-1930*", *American Journal of Sociology* 38: 71-78.

BELL, D. (1964): "Twelve modes of prediction-a preliminary sorting of approaches to the social sciences", *Daedalus* 93, 3: 845-880.

BERELSON, B. (1952): *Content Analysis in Communications Research*, Nueva York, Free Press.

—— y P.F. LAZARSFELD (1948): *The Analysis of Communication Content*, Chicago y Nueva York, University of Chicago and Columbia University.

BERELSON, B. y P.J. SALTER (1946): "Majority and minority Americans: an analysis of magazine fiction", *Public Opinion Quarterly* 10: 168-190.

BERELSON, B. y G.A. STEINER (1964): *Human Behavior: An Inventory of Scientific Findings*, Nueva York, Harcourt Brace Jovanovich.

BERKMAN, D. (1963): "Advertising in *Ebony* and *Life*: Negro aspirations vs. reality", *Journalism Quarterly* 40: 53-64.

BRODER, D.P. (1940): "The adjective-verb quotient: a contribution to the psychology of language", *Psychological Record* 3: 310-343.

BROOM, L. y S. REECE (1955): "Political and racial interest: a study in content analysis", *Public Opinion Quarterly* 19: 5-19.

BROUWER, M., C.C. CLARK, G. GERBNER y K. KRIPPENDORFF (1969): "The television world of violence", págs. 311-339 y 519-591, en R.K. Baker y S.J. Ball (comps.), *Mass Media and Violence: A Report to the National Commission*

on the Causes and Prevention of Violence, Washington, DC, Government Printing Office.

BUDD, R.W. (1964): "Attention score: a device for measuring news 'play' ", *Journalism Quarterly* 41: 259-262.

CAHNMAN, W.J.(1948): "A note on marriage announcements in the *New York Times*", *American Sociological Review* 13: 96-97.

CAMPBELL, D.T. (1957): "Factors relevant to the validity of experiments in social settings", *Psychological Bulletin* 54, 4: 297-311.

— y D.W. FISKE (1959): "Convergent and discriminant validation by the multitrait-multimethod matrix", *Psychological Bulletin* 56, 2: 81-105.

CANTRIL, H.(1948): "Opinion trends in World War II: some guides to interpretation", *Public Opinion Quarterly* 12: 30-44.

CARTWRIGHT, D.P. (1953): "Analysis of qualitative material", págs. 421-470, en L. Festinger y D. Katz (comps.), *Research Methods in the Behavioral Sciences*, Nueva York, Holt, Rinehart & Winston.

CHOMSKY, N.(1959): "Review of B.F. Skinner, *Verbal Behavior*", *Language* 35: 26-58.

COHEN, B.C. (1957): *The Political Process and Foreign Policy: The Making of the Japanese Peace Settlement*, Princeton, NJ, Princeton University Press.

COHEN, J. (1960): "A coefficient of agreement for nominal scales", *Educational and Psychological Measurement* 20, 1: 37-46.

"Content analysis: a new evidentiary technique" (1948), *University of Chicago Law Review* 15: 910-925.

Council on Inter-Racial Books for Children (1977): *Stereotypes, Distortions and Ommissions in U.S. History Textbooks*, Nueva York, Racism and Sexism Resource Center for Educators.

DALE, E.(1937): "The need for the study of newsreels", *Public Opinion Quarterly* 1: 122-125.

DEICHSEL, A. (1975): *Elektronische Inhaltsanalyse*, Berlín, Volker Spiess.

De WEESE, L.C., III (1977): "Computer content analysis of 'day-old' newspapers: a feasibility study", *Public Opinion Quarterly* 41: 91-94.

DIBBLE, V.K.(1963): "Four types of inferences from documents to events", *History and Theory* 3, 2: 203-221.

DOLLARD, J.y F. AULD, (h.) (1959): *Scoring Human Motives: A Manual*, New Haven, CT, Yale University Press.

DOLLARD, J. y O.H. MOWRER (1947): "A method of measuring tension in written documents", *Journal of Abnormal and Social Psychology* 42: 3-32.

DOVRING, K. (1954-1955): "Quantitative semantics in 18th century Sweden", *Public Opinion Quarterly* 18, 4: 389-394.

DUNPHY, D.C. (1966): "The construction of categories of content analysis dictionaries", págs. 134-168, en P.J. Stone y otros (comps.), *The General Inquirer*, Cambridge, MIT Press.

DZIURZYNSKY, P.S. (1977): "Development of a content analytic instrument for advertising appeals used in prime time television commercials", Tesis de doctorado, Universidad de Pennsylvania.

EKMAN, P. y W.V. FRIESEN(1968): "Nonverbal behavior in psychotherapy research", en J. Shlien (comp.), *Research in Psychotherapy*, Washington, DC, American Psychological Association.

— y T.G. TAUSSIG (1969): "VID-R and SCAN: tools and methods for the auto-

mated analysis of visual records", págs. 297-312, en G.Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

ELLISON, J.W. (1965): "Computers and the testaments", págs. 64-74, en *Proceedings, Conference on Computers for the Humanities*, New Haven, CT, Yale University Press.

ELMAGHRABY, S.E. (1970): "Theory of networks and management science". *Management Science* 17: 1-34, 1354-1371.

ERTEL, S. (1976): "Dogmatism: an approach to personality", págs. 34-44, en A. Deichsel y K. Holzenscheck (comps.), *Maschinelle Inhaltsanalyse, Materialen I.*, Hamburgo, University of Hamburg.

FENTON, F. (1910): "The influence of newspaper presentations on the growth of crime and other anti-social activity", *American Journal of Sociology* 16: 342-371, 538-564.

FLESH, R. (1948): "A new readability yardstick", *Journal of Applied Psychology* 32: 221-233.

— (1951): *How To Test Readability*, Nueva York, Harper & Row.

FOSTER, C.R. (1938): *Editorial Treatment of Education in the American Press*, Cambridge, MA, Harvard University Press.

FOSTER, H.S. (1935): "How America became belligerent: a quantitative study of war news 1914-17", *American Journal of Sociology* 40: 464-475.

GARFIELD, E. (1979): *Citation Index: Its Theory and Application to Science, Technology and Humanities*, Nueva York, John Wiley.

GELLER, A., D. KAPLAN y H.D. LASSWELL (1942): "An experimental comparison of four ways of coding editorial content", *Journalism Quarterly* 19: 362-370.

GEORGE, A.L. (1959a): *Propaganda Analysis: A Study of Inferences Made from Nazi Propaganda in World War II*, Evanston, IL, Row. Peterson.

— (1959b): "Quantitative and qualitative approaches to content analysis", págs. 7-32, en I. de Sola Pool (comp.), *Trends in Content Analysis*, Urbana, University of Illinois Press.

GERBNER, G. (1959): "The social role of the confession magazine", *Social Problems*, 6: 29-40.

— (1964): "Ideological perspectives and political tendencies in news reporting", *Journalism Quarterly* 41: 495-508.

— (1966): "An institutional approach to mass communications research", págs. 429-445, en L. Thayer (comp.), *Communication: Theory and Research*, Springfield, IL, Charles C. Thomas.

— (1969): "Toward 'cultural indicators': the analysis of mass mediated public message systems", págs. 123-132, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

— y G. MARVANJI (1977): "The many worlds of the world's press", *Journal of Communication* 27, 1: 52-75.

GERBNER, G., L. GROSS, N. SIGNORIELLI, M. MORGAN Y M. JACKSON-BEECK (1979): "Violence profile no. 10: trends in network television drama and view conceptions of social reality, 1967-1978", Annenberg School of Communications, University of Pennsylvania (mimeografiado).

GERBNER, G., O.R. HOLSTI, K. KRIPPENDORFF, W.J. PAISLEY y P.J. STONE (comps.) (1969): *The Analysis of Communication Content: Developments in Scientific Theories and Computer Techniques*, Nueva York, John Wiley.

GIEBER, W. (1964): "News in what newspaperment make it", págs. 173-182, en L.A. Dexter y D.M. White (comps.), *People, Society and Mass Communications*, Nueva York, Free Press.

GREEN, B.F., A.K. WOLF, C. CHOMSKY y K. LAUGHERY (1963): "Baseball: an automatic question answerer", págs. 207-216, en E.A. Feigenbaum y J. Feldman (comps.), *Computers and thought*, Nueva York, MacGraw-Hill.

GROTH, O. (1948): *Die Geschichte der Deutschen Zeitungswissenschaft, Probleme und Methoden*, Munich, Konrad Weinmayer.

HACHT, D.L. y M. HACHT (1947): "Criteria of social status as derived from marriage announcements in the *New York Times*", *American Sociological Review* 12: 396-403.

HAYS, D.C. (1960): *Automatic Content Analysis*, Santa Monica, CA, Rand Corporation.

— (1969): "Linguistic foundations for a theory of content analysis", págs. 57-67, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

HERDAN, G. (1960): *Type-Token Mathematics: A Textbook of Mathematical Linguistics*, La Haya, Mouton.

HERMA, H., E. KRISS y J. SHOR (1943): "Freud's theory of the dream in American textbooks", *Journal of Abnormal and Social Psychology* 38: 319-334.

HOLSTI, O.R. (1962): "The belief system and national images: John Foster Dulles and the Soviet Union", Tesis de doctorado, Stanford University.

— (1969): *Content Analysis for the Social Sciences and Humanities*, Reading, MA, Addison-Wesley.

— R.A. BRODY y R.C. NORTH (1965): "Measuring affect and action in international reaction models: empirical materials from the 1962 Cuban crisis", *Peace Research Society Papers* 2: 170-190.

IKER, H.P. (1975): *WORDS System Manual*, Rochester, NY, Computer Printout.

— (s.f.) "SELLECT: a computer program to identify associationally rich worlds for content analysis", Nueva York (mimeografiado).

INNIS, H.A. (1951): *The Bias of Communication*, Toronto, University of Toronto Press.

Institute for Propaganda Analysis, Inc. (1937): "How to detect propaganda", *Propaganda Analysis* 1: 5-8.

JANDA, K. (1969): "A microfilm and computer system for analysing comparative politics literature", págs. 407-435, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

JANIS, I.L. (1965): "The problem of validating content analysis", págs. 55-82, en H.D. Lasswell y otros (comps.), *Language of Politics*, Cambridge, MIT Press.

— y R.H. Fadner (1965): "The coefficient of imbalance", págs. 153-169, en H.D. Lasswell y otros (comps.), *Language of Politics*, Cambridge, MIT Press.

KAPLAN, A. y J.M. GOLDSEN (1965): "The reliability of content analysis categories", págs. 83-112, en H.D. Lasswell y otros (comps.), *Language of Politics*, Cambridge, MIT Press.

KATZ, E., M. GUREVITCH, B. DANET y T. PELED (1967): "Petitions and prayers: a content analysis of persuasive appeals", Univesity of Chicago (mimeografiado).

KLAUSNER, S.Z. (1968): "Two centuries of child-rearing manuals", Informe técnico a la Joint Commission on Mental Health of Children, Department of Sociology, University of Pennsylvania (mimeografiado).

KLEIN, M.W. y N. MACCOBY (1954): "Newspaper objectivity in the 1952 campaign", *Journalism Quarterly* 31: 285-296.

KLEIST, K. (1934): *Gehirn Pathologie*, Leipzig, Johan Ambrosius Barth.

KLIR, G. y M. VALACH (1965): "Language as means of communication between man and machine", págs. 315-373, en *Cybernetic Modelling*, Princeton, NJ, D. Van Nostrand.

KOPPELAAR, H., B. VAN KONIGSVELD y H. WEIJENBURG (1978): "Verbale modelvorming", *Sociologische Gids* 25, 3: 202-211.

KOPS, M. (1977): *Answahlverfahren in der Inhaltsanalyse*, Meisenheim/Glan, Anton Hain.

KRACAUER, S. (1947): *From Caligari to Hitler: A Psychological History of German Film*, Londres, Dennis Dobson. [Trad. esp.: *De Caligari a Hitler*, Barcelona, Paidós, 1985.]

— (1952-1953): "The challenge of quantitative content analysis", *Public Opinion Quarterly* 16: 631-642.

KRENDEL, E.S. (1970): "A case study of citizen complaints as social indicators", *IEEE Transactions on Systems Science and Cybernetics*, SSC-6, 4: 267-272.

KRIPPENDORFF, K. (1969a): "Theories and analytical constructs: introduction", págs. 3-16, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

— (1969b): "Models of messages: three prototypes", págs. 69-106, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

— (1970a): "On generating data in communication research", *Journal of Communication*, 20, 3: 241-269.

— (1970b): "The expression of value in political documents", *Journalism Quarterly* 47: 510-518.

— (1971): "Reliability of recording instructions: multivariate agreement for nominal data", *Behavioral Science* 16, 3: 222-235.

— (1974): "An algorithm for simplifying the representation of complex systems", págs. 1693-1702, en J. Rose (comp.), *Advances in Cybernetics and Systems*, Nueva York, Gordon & Breach.

— (1980a): "Validity in content analysis", págs. 69-112, en E. Mochmann (comp.), *Computerstrategien fuer die Kommunikations analyse*, Frankfurt, Campus.

— (1980b): "Clustering", en P.R. Monge y J.N. Capella (comps.), *Multivariate Techniques in Communication Research*, Nueva York, Academic Press.

LABOV, W. (1972): *Sociolinguistic Patterns*, Philadelphia, University of Pennsylvania Press.

LANDIS, H.H. y H.E. BURTT (1924): "A study of conversations", *Journal of Comparative Psychology* 4: 81-89.

LASSWELL, H.D. (1927): *Propaganda Technique in the World War*, Nueva York, Knopf.

— (1938): "A provisional classification of symbol data", *Psychiatry* 1: 197-204.

— (1941): "The world attention survey: an exploration of the possibilities of

studying attention being given to the United States by newspapers abroad", *Public Opinion Quarterly* 5, 3: 456-462.

— (1960): "The structure and function of communication in society", págs. 117-130, en W. Schramm (comp.), *Mass Communications*, Urbana, University of Illinois Press.

— (1963): *Politics: Who Gets What, When, How*, Nueva York, Meridian.

— (1965a): "Why be quantitative?", págs. 40-52, en H.D. Lasswell y otros (comps.), *Language of Politics*, Cambridge, MIT Press.

— (1965b): "Detection: propaganda detection and the courts", págs. 173-232, en H.D. Lasswell y otros (comps.), *Language of Politics*, Cambridge, MIT Press.

— y A. KAPLAN (1950): *Power and Society: A Framework for Political Inquiry*, New Haven, CT, Yale University Press.

LASSWELL, H.D., D. LERNER e L DE SOLA POOL (1952): *The Comparative Study of Symbols*, Stanford, CA, Stanford University Press.

LASSWELL, H.D. y otros (1965): *Language of Politics: Studies in Quantitative Semantics*, Cambridge, MIT Press.

LAZARSFELD, P.F., B. BERELSON y H. GAUDET (1948): *The People's Choice: How the Voter Makes Up His Mind in a Presidential Campaign*, Nueva York, Columbia University Press.

LEITES, N., E. BERNAUT y R.L. GARTHOFF (1951): "Politburo images of Stalin", *World Politics* 3: 317-339.

LINDSAY, R.K. (1963): "Inferential memory as the basis of machines which understand natural language", pág. 217-233, en E.A. Feigenbaum y J. Feldman (comps.), *Computers and Thought*, Nueva York, McGraw-Hill.

LIPPMANN, W. (1922): *Public Opinion*, Nueva York, Macmillan.

LOEBL, E. (1903): *Kultur und Presse*, Leipzig, Duncker & Humbolt.

LOEVENTHAL, L. (1944): "Biographies in popular magazines", págs. 507-548, en P.F. Lazarsfeld y F.N. Stanton (comps.), *Radio Research* 1942-1943, Nueva York, Deuell, Sloan & Pearce.

LORR, M. y D.M. MCNAIR (1966): "Methods relating to evaluation of therapeutic outcome", págs. 573-594, en L.A. Gottschalk y A.H. Auerbach (comps.), *Methods of Research in Psychotherapy*, Englewood Cliffs, NJ, Prentice-Hall.

LINCH, K. (1965): *The Image of the City*, Cambridge, MIT Press.

MACCOBY, N., F.O. SABGHIR y B. CUSHING (1950): "A method for the analysis of news coverage of industry", *Public Opinion Quarterly* 14: 753-758.

MAHL, G.F. (1959): "Exploring emotional states by content analysis", págs. 89-130, en I. de Sola Pool (comp.), *Trends in Content Analysis*, Urbana, University of Illinois Press.

MANN, M.B. (1944): "Studies in language behavior: III. The quantitative differentiation of samples of written language". *Psychological Monographs* 56, 2: 41-47.

MARANDA, P. y E.K. MARANDA (comps.) (1971): *Structural Analysis of Oral Tradition*, Filadelfia, University of Pennsylvania Press.

MARKOV, A.A. (1913): "Essai d'une recherche statistique sur le texte du roman 'Eugène Onegin' illustrant la liaison des épreuves en chaîne (ruso)", *Bulletin de L'Académie Impériale des Sciences de St. Pétersbourg* 6, 7: 153-162.

MARTIN, H. (1936): "Nationalism and children's literature", *Library Quarterly* 6: 405-418.

MATTHEWS, B.C. (1910): "A study of a New York daily", *Independent* 68: 82-86.

MCCLELLAND, D.C. (1958): "The use of measures of human motivation in the study of society", págs. 518-552, en J.W. Atkinson (comp.), *Motives in Fantasy, Action and Society*, Princeton, NJ, D. Van Nostrand.

MCDIARMID, J. (1937): "Presidential inaugural addresses: a study in verbal symbols", *Public Opinion Quarterly* 1: 79-82.

MERRILL, J.C. (1962): "The image of the United States in ten Mexican dailies", *Journalism Quarterly* 39: 203-209.

MERRITT, R.L. (1966): *Symbols of American Community. 1735-1775*, New Haven. CT, Yale University Press.

MIDDLETON, R. (1960): "Fertility values in American magazine fiction: 1916-1956", *Public Opinion Quarterly* 24: 139-143.

MILES, J. (1951): "The continuity of English poetic language", págs. 517-535, en *University of California Publications in English*, Berkeley, University of California Press.

MILLER, G.A. (1951): *Language and Communication*, Nueva York, McGraw-Hill.

MINSKY, M. (comp.) (1968): *Semantic Information Processing*. Cambridge, MIT Press.

MORTON, A.Q. (1963): "A computer challenges the church", *Observer* (noviembre 3).

— y M. LEVINSON (1966): "Some indications of authorship in Green prose". págs. 141-179, en J. Leed (comp.), *The Computer and Literary Style*, Kent, OH, Kent State University Press.

MOSTELLER, F. y D.L. WALLACE (1963): "Inference in an authorship problem", *Journal of American Statistical Association* 58: 275-309.

— (1964): *Inference and Disputed Authorship: The Federalist*, Reading, MA., Addison-Wesley.

MURRAY, E.J., F. AULD (h.) y A.M. WHITE (1954): "A psychotherapy case showing progress but no decrease in the discomfort-relief quotient", *Journal of Consulting Psychology* 18: 349-353.

NAMENWIRTH, J.Z. (1973): "Wheels of time and the interdependence of value change in America", *Journal of Interdisciplinary History* 3: 649-683.

NEWELL, A. y H.A. SIMON (1956): "The logic theory machine", *IRE-Transactions on Information Theory* IT-2 3: 61-79.

— (1963): "General Problem Solver: a program that simulates human thought", págs. 279-293, en E.A. Feigenbaum y J. Feldman (comps.), *Computers and Thought*, Nueva York, McGraw-Hill.

NIXON, R.B. y R.L. JONES (1956): "The content of non-competitive newspapers", *Journalism Quarterly* 33: 299-314.

NORTH, R.C., O.R. HOLSTI, M.G. ZANINOVICH y D.A. ZINNES (1963): *Content Analysis: A Handbook with Applications for the Study of International Crisis*, Evanston, IL, Northwester University.

OSGOOD, C.E. (1959): "The representation model and relevant research methods", págs. 33-88, en I. de Sola Pool (comp.), *Trends in Content Analysis*, Urbana, University of Illinois Press.

— S. SPORTA y J.C. NUNNALLY (1956): "Evaluative assertion analysis", *Litera* 3: 47-102.

—— G.J. SUCI y P.H. TANNENBAUM (1957): *The Measurement of Meaning*, Urbana, University of Illinois Press.

O'SULLIVAN, T.C. (h.) (1961): "Factor analysis concepts identified in theoretical writtings—an experiment design", *Itek Laboratories*, Lexington, MA. (septiembre).

PAISLEY, W.J. (1964): "Identifying the unknown communicator in painting, literature and music: the significance of minor encoding habits", *Journal of Communication* 14: 219-237.

PHILLIPS, D.P. (1978): "Airplane accident fatalities increase just after newspaper stories about murder and suicide", *Science* 201: 748-749.

PIAULT, C. (1965): "A methodological investigation of content analysis using electronic computers for data processing", págs. 273-293, en D. Hymes (comp.), *The Use of Computers in Anthropology*, La Haya, Mouton.

PIERCE, B. (1930): *Civic Attitudes in American School Textbooks*, Chicago, University of Chicago Press.

POOL, I. DE SOLA (1951): *Symbols of Internationalism*, Stanford, CA, Stanford University Press.

—— (1952a): *The "Prestige Papers": A Survey of Their Editorials*, Stanford, CA, Stanford University Press.

—— (1952b): *Symbols of Democracy*, Stanford, CA, Stanford University Press.

—— (comp.) (1959): *Trends in Content Analysis*, Urbana, University of Illinois Press.

—— R.P. ABELSON y S.L. POPKIN (1964): *Candidates, Issues and Strategies: A Computer Simulation of the 1960 Presidential Election*, Cambridge, MIT Press.

RAINOFF, T.J. (1929): "Wave-like fluctuations of creative productivity in the development of West-European physics in the eighteenth and nineteenth centuries", *ISIS* 12: 287-307.

RAPOPORT, A. (1969): "A system-theoretic view of content analysis", págs. 17-38, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

REYNOLDS, H.T. (1977): *Analysis of Nominal Data*, Beverly Hills, CA, Sage.

Roget's International Thesaurus (1974): Londres, Collins.

SCHANK, R.C. y R.P. ABELSON (1977): *Scripts, Plans, Goals and Understanding: An Inquiry into Human Knowledge Structures*, Hillsdale, NJ, Erlbaum. [Trad. esp.: *Guiones, planes, metas y entendimiento*, Barcelona, Paidós, 1989.]

SCHUTZ, W.C. (1958): "On categorizing qualitative data in content analysis", *Public Opinion Quarterly* 22, 4: 503-515.

SCOTT, W.A. (1955): "Reliability of content analysis: the case of nominal scale coding", *Public Opinion Quarterly* 19: 321-325.

SEBEOK, T.A. y L.H. ORZACK (1953): "The structure and content of cheremis charms", *Anthropos* 48:369-388.

SEBEOK, T.A. y V.J. ZEPS (1958): "An analysis of structured content with application of electronic computer research in psycholinguistics", *Language and Speech* 1: 181-193.

SEDELOW, S.Y. (1967): *Stylistic Analysis*, Santa Monica, CA., SDC.

—— y W.A. SEDELOW(h.) (1966): "A preface to computational stylistics", en J.

Leed (comp.), *The Computer and Literary Style*, Kent, OH, Kent State University Press.

SELLTIZ, C., M. JAHODA, M. DEUTSCH y S.W. COOK (1964): *Research Methods in Social Relations*, Nueva York, Holt, Rinehart & Winston.

SHANAS, E. (1945): "The *American Journal of Sociology* through fifty years", *American Journal of Sociology* 50: 522-533.

SHANNON, C.E. y W. WEAVER (1949): *The Mathematical Theory of Communication*, Urbana, University of Illinois Press.

SHNEIDMAN, E.S. (1963): "The logic of politics", págs. 178-199, en L. Arons y M.A. May (comps.), *Television and Human Behavior*, Englewood Cliffs, NJ, Prentice-Hall.

—— (1966): *The Logics of Communication: A Manual for Analysis*, China Lake, CA, U.S. Naval Ordnance Test Station.

—— (1969): "Logical content analysis: an exploration of styles of concludifying", págs. 261-279, en G. Gerbner y otros (comps.), *The Analysis of Communication Content*, Nueva York, John Wiley.

SIMPSON, G.E. (1934): "The Negro in the Philadelphia press", Tesis de doctorado, University of Pennsylvania.

SINGER, J.D. (1964): "Media analysis in inspection for disarmament", *Journal of Arms Control* 1: 248-260.

SMYTHE, D.W. (1954): "Some observations on communications theory", *Audio-Visual-Communication Review* 2, 1: 248-260.

SPEED, G.J. (1893): "Do newspapers now give the news?", *Forum* 15: 705-711.

SPIEGELMAN, M., C. TERWILLIGER y F. FEARING (1953a): "The reliability of agreement in content analysis", *Journal of Social Psychology* 37: 175-187.

—— (1953b): "The content of comics: goals and means to goals of comic strip characters", *Journal of Social Psychology* 37: 189-203.

STEMPEL, G.H., III (1952): "Sample size for classifying subject matter in dailies: research in brief", *Journalism Quarterly* 29: 333-334.

STEVENS, S.S. (1946): "On the theory of scales of measurement", *Science* 103, 2684: 677-680.

STONE, P.J. y E.B. HUNT (1963): "A computer approach to content analysis using the general inquirer system", págs. 241-256, en E.C. Johnson (comp.), *American Federation of Information Processing Societies, Conference Proceedings*, Baltimore, Autor.

STONE, P.J., D.C. DUNPHY, M.S. SMITH y D.M. OGILVIE (1966): *The General Inquirer: A Computer Approach to Content Analysis*, Cambridge, MIT Press.

STREET, A.T. (1909): "The truth about newspapers", *Chicago Tribune* (julio 25).

SUCI, G. y T.R. HUSEK (1957): "A content analysis of 70 messages with 55 variables and subsequent effect studies", Institute for Communication Research, University of Illinois, Urbana (mimeografiado).

TANNENBAUM, P.H. y B.S. GREENBERG (1961): "*J.Q.* references: a study of professional change", *Journalism Quarterly* 38: 203-207.

TAYLOR, W.L. (1953): " 'Cloze procedure': a new tool for measuring readability", *Journalism Quarterly* 30: 415-433.

TENNEY, A.A. (1912): "The scientific analysis of the press", *Independent* 73: 895-898.

THOMPSON, S. (1932): *Motif-Index of Folk Literature: A Classification of*

Narrative Elements in Folk-Tales, Ballads, Myths, Fables, Mediaeval Romances, Exempla, Fabliaux, Jest-Books, and Local Legends, Bloomingdale, Indiana University Studies.

TUKEY, J.W. (1980): "Methodological comments focused on opportunities", en P.R. Monge y J.N. Cappella (comps.), *Multivariate Techniques in Communication Research*, Nueva York, Academic Press.

WALWORTH, A. (1938): *Social Histories at War: A Study of the Treatment of Our Wars in the Secondary School History Books of the United States and in Those of Its Former Enemies*, Cambridge, MA, Harvard University Press.

WATZLAWICK, P. J.H. BEAVIN y D.D. JACKSON (1967): *Pragmatics of Human Communication*, Nueva York, W.W. Norton.

WEBB, E.J., D.T. CAMPBELL y R.D. SCHWARTZ (1966): *Unobtrusive Measures: Nonreactive Research in the Social Sciences*, Chicago, Rand McNally.

WEBER, M. (1911): " 'Geschaeftsbericht' in Verhandlungen des ersten deutschen Soziologietages vom 19.-22. Oktober 1910 in Frankfurt a. M.", págs. 39-62, en *Schrift der Deutschen Gesellschaft fuer Soziologie*.

Webster's New Dictionary of Symbols (1973), Springfield, MA, Merriam.

Webster's Seventh New Collegiate Dictionary (1967), Springfield, MA, Merriam.

WEIK, K.E. (1968): "Systematic observational methods", págs. 357-451, en G. Lindzey y E. Aronson (comps.), *The Handbook of Social Psychology*, Reading, MA, Addison-Wesley.

WHITE, D.M. (1964): "The 'Gatekeeper': a case study in selection of news", págs. 160-172, en L.A. Dexter y D.M. White (comps.), *People, Society and Mass Communication*, Nueva York, Free Press.

WHITE, P.W. (1924): "Quarter century survey of press content shows demand for facts", *Editor and Publisher*, 57, mayo 31.

WHITE, R.K. (1947): "*Black Boy*, a value analysis", *Journal of Abnormal and Social Psychology* 42: 440-461.

— (1951): *Value-Analysis: The Nature and Use of the Method*, Glen Gardiner, NJ, Libertarian Press.

WILCOX, D.F. (1900): "The American newspaper: a study in social psychology", *Annals of American Academy of Political and Social Science* 16: 56-92.

WILLEY, M.M. (1926): *The Country Newspaper: A Study of Socialization and Newspaper Content*, Chapel Hill, University of North Carolina Press.

WOODWARD, J.L. (1934): "Quantitative newspaper analysis as a technique of opinion research", *Social Forces* 12: 526-537.

WRIGHT, C.E. (1964): "Functional analysis and mass communication", págs. 91-109, en L.A. Dexter y D.M. White (comps.), *People, Society and Mass Communication*, Nueva York, Free Press.

YULE, G.U. (1944): *The Statistical Study of Literary Vocabulary*, Londres, Cambridge University Press.

ZADEH, L.A. (1965): "Fuzzy sets", *Information and Control* 8, 3: 338-353.

ZILLMAN, D.(1964): *Konzept der Semantischen Aspektanalyse*, Zurich, Institut fuer Kommunikationsforschung.

ZUCKER, H.G. (1978): "The variable nature of news media influence", *Communication Yearbook* 2: 225-240.